

STIC-Biotech/Ch mLib

79027

Fr m: Mehta, Ashwin
Sent: Wednesday, October 30, 2002 9:06 AM
To: STIC-Biotech/ChemLib
Subject: sequence search

STIC,

Please search the commercial and interference databases for the following from 09/896,186

- 1) the nucleotide sequence of SEQ ID NO: 23
- 2) the amino acid sequence of SEQ ID NO: 24

My mail room is 9E12, office 9E07, art unit 1638.

Thank you,
Ashwin

Ashwin Mehta
United States Patent and Trademark Office
Biotechnology Patent Examiner
703-306-4540

RECEIVED
OCT 30 2002
STIC

Point of Contact
P. Sheppard

Searcher's telephone number: (703) 308-4499
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: 11/7/02
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 2, 2002, 14:04:54 ; Search time 2044 Seconds
(without alignments)
8311.676 Million cell updates/sec

Title: US-09-896-186B-23

Perfect score: 1049

Sequence: 1 accaaagcattatatttatt.....tttaaccgtccagaactag 1049

Scoring table: IDENTITY_NUC
Gapop 10.0 ; Gapext 1.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estcov:*
6: em_estro:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	288.8	27.5	435	9	AU226180 AU226180
C 2	273.8	26.1	840	10	BE659005 BE659005
C 3	261	24.9	704	12	BG41595 GA_Ea001
C 4	245	23.4	552	10	AM201789 AM201789
C 5	243.8	23.2	671	9	AU237507 AU237507
C 6	241.2	23.0	467	10	AM832139 AM832139

7	235.4	22.4	661	14	BO799068	BO799068 EST 1237
8	233.8	22.3	535	13	BM406124	BM406124 EST580451
9	223	21.3	506	12	BG133320	BG133320 EST466112
10	213.6	20.4	531	13	BM094501	BM094501 saj16903
11	172.2	16.4	515	13	BM358892	BM358892 GA_Ea001
12	169.6	16.2	517	13	BM358891	BM358891 GA_Ea001
13	167.4	16.0	555	14	BM731125	BM731125 sa168a05
14	162	15.4	595	13	Bu213300	Bu213300 Bu213300
15	156.8	14.9	499	13	BI433038	BI433038 EST535799
16	156.6	14.9	556	14	BM890816	BM890816 sam08c02
17	156	14.9	563	13	BM527093	BM527093 sa149e10
18	152	14.5	531	10	AM460204	AM460204 sh56c10.y
19	151	14.4	684	12	BE589996	BE589996 EST497838
20	143.2	13.7	625	12	BE999775	BE999775 EST431498
21	142.4	13.6	560	12	BG157629	BG157629 saa88c03
22	140.8	13.4	572	9	AU235472	AU235472 AU235472
23	134.4	12.8	680	13	Bu321563	Bu321563 Bu321563
24	129.6	12.4	407	12	BE821184	BE821184 GW700024A
25	129	12.3	220	17	BH849559	BH849559 SALK_0698
26	124.6	11.9	431	9	A1748378	A1748378 sb52a06.y
27	117	11.2	445	10	AM119478	AM119478 sda6a06.y
28	100.6	9.6	494	10	AM720170	AM720170 LjNEST16b
29	99	9.4	654	13	Bu220795	Bu220795 Bu220795
30	97.8	9.3	660	17	B61114	B61114 T20KRF TAM
31	96	9.2	600	13	Bu316072	Bu316072 Bu316072
32	84.6	8.1	550	14	BM785803	BM785803 K-EST0064
33	76.2	7.3	405	17	CNS00V8H	AL091677 Arabidops
34	66.8	6.4	533	17	AO580338	AO580338 T135971b
35	65.4	6.2	473	17	AO866027	AO866027 nbe00026P
36	65.4	6.2	901	17	AO867899	AO867899 nbe00018H
37	62	5.9	342	17	AO841745	AO841745 T132158b
38	62	5.9	457	17	BH168898	BH168898 SALK_0003
39	57.2	5.5	447	13	BI420478	BI420478 LjNEST57f
40	56.6	5.4	586	13	BM191106	BM191106 da32b10
41	55.6	5.3	569	17	AO580359	AO580359 T136000b
42	50.2	4.8	151	10	BE022823	BE022823 sm89c11.y
43	47	4.5	711	13	Bu169253	Bu169253 Bu169253
44	46.6	4.4	620	9	AL647838	AL647838 AL647838
45	45.8	4.4	161	17	B24158	B24158 F19B10TR IG

ALIGNMENTS

RESULT 1
AU226180/c
LOCUS
DEFINITION
AU226180 RAFL14 Arabidopsis thaliana cDNA clone RAFL14-14-A20 3',
mRNA sequence.
ACCESSION
AU226180
VERSION
AU226180.1 GI:19740827
KEYWORDS
EST.
SOURCE
Arabidopsis thaliana
ORGANISM
thale cress.
REFERENCE
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Akiyama, K., Enju, A., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Atakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y. and Shinozaki, K.
Large scale analysis of Arabidopsis full-length cDNA
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@frc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI

and XhoI was ligated to modified Lambda FL-C1 vector (Carinici et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified Bluescript vector. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

FEATURES

source

```

1.435
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone_lib="RAFL14-14-820"
/clone_lib="RAFL14"
/tissue_type="root"
/lab_host="DH10B"
/note="Site_1: BamHI; Site_2: SalI"
BASE COUNT      117 a      85 c      85 g      148 t
ORIGIN

```

Query Match

Best Local Similarity 90.3%; Pred. No. 2,2e-61; Matches 327; Conservative 0; Mismatches 17; Indels 18; Gaps 1;

```

QY 706 GAGATAAAAATGGGCGCTGGCTCACTAAGACACATGTTGGCAAGAGCTCCTGA 765
    |||||||
Db 435 GAGATAAAAATGGGCGCTGGCTCACTAAGACACATGTTGGCAAGAGCTCCTGA 376
QY 766 AGCCAAACAGAAATCAGGCTTGGGAACTGGAGTTTATCCTCTGTCAAGACAGCTTAC 825
    |||||||
Db 375 AGCCAAACAGAAATCAGGCTTGGGAACTGGAGTTTATCCTCTGTCAAGACAGCTTAC 316
QY 826 AATAGCAGCAGCAGGATGCTTATGCTTCAATGCGATCTTTAACAAGTTCTTAAGACCTTC 885
    |||||||
Db 315 AATAGCAGCAGCAGGATGCTTATGCTTCAATGCGATCTTTAACAAGTTCTTAAGACCTTC 256
QY 886 CTGATGCTGTCAGTGCCTCACTAAGTGAAGAGAGAGCTTAAAGTTAGGCTTAAACCCC 945
    |||||||
Db 255 CTGATGCTGTCAGTGCCTCACTAAGTGAAGAGAGAGCTTAAAGTTAGGCTTAAACCCC 196
QY 946 AAGATTAGCAGCAATATGATATGATACACCTAATCTAGTCAAGTATGCAATTTCTT 1002
    |||||||
Db 195 AAGATTAGCAGCAATATGATATGATACACCTAATCTAGTCAAGTATGCAATTTCTTATG 136
QY 1003 -----GTGATATTTGATCTAGTCTGCTGTCGCCCTTAAACGCTCCAGAAACT 1047
    |||||
Db 135 GGAATTACAAAGATATTGAACACAGATATTAGTGGCGTCCATTAATCTTCCAGAAACT 76
QY 1048 AG 1049
    ||
Db 75 AG 74

```

RESULT 2
BE659005/c 840 bp mRNA linear EST 24-MAY-2001
LOCUS GM700008A10G11 Gm-r1070 Glycine max cDNA clone Gm-r1070-2997 3',
DEFINITION mRNA sequence.
ACCESSION BE659005
VERSION BE659005.1 GI:9984897
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 840)

Vodkin, L., Keim, P., Shoemaker, R., Retzel, E., Khanna, A., Coryell, V.,
Erpelding, J., Rapp, C., Shoop, E., Pardinas, J., Liu, L. and Lewin, H.,
A Functional Genomics Program for Soybean (NSF 9872565)
Unpublished (1999)
Other ESTs: AW460204 corresponding to Gm-cl015-3787 (5')
Contact: Vodkin, L.O., PI, A Functional Genomics Program for
Soybean (NSF 9872565)

Lewin, H. A., Director, Keck Center for Comparative and Functional

Genomics
University of Illinois
Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
Tel: (217) 244-6147
Fax: (217) 333-4582
Email: l-vodkin@uiuc.edu
This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134. For further information
call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or info@genome
systems.com web site: www.genomesystems.com
Seq primer: 5'-TTTTTTTTTTTTTTTTT(A/C/G)-3'.

FEATURES

source

```

1.840
/organism="Glycine max"
/db_xref="taxon:3847"
/clone_lib="Gm-r1070-2997"
/clone_lib="Gm-r1070"
/note="The library Gm-r1070 is a sequence-driven, rerecked
set of 9,216 clones selected from cDNA libraries from
various tissues and stages of development of soybean that
represent 2,639 sequences from immature cotyledons, 1,770
from immature seed coats, 3,938 from flowers, and 869
from young pods. The 5' ESTs of the source clones from
the different libraries was used to select singletons, or
a representative of each contig, which were rerecked to
form library Gm-r1070. The cDNA clones of the rerecked
Gm-r1070 library were then sequenced at the 3' end. The
contig analysis to select unique genes was performed by
the Laboratory of Ernest Retzel, Center for Computational
Genomics and Bioinformatics, University of Minnesota,
http://www.cbc.umn.edu/ResearchProjects/soybean/index.html
. Rerecking was performed by Genome Systems, St. Louis,
http://www.genomesystems.com, and 3' sequencing by the
Keck Center for Comparative and Functional Genomics,
University of Illinois,
http://www.life.uiuc.edu/biotech/keck.html. Note: The
corresponding 5' EST from each clone in the Gm-r1070
library is listed in the 'OTHER EST' field. The detailed
information on the source library for each clone can also
be obtained by referring to the Genome Systems clone ID of
the original cDNA library that is also listed under
'OTHER EST'."
BASE COUNT      242 a      167 c      158 g      240 t      33 others
ORIGIN

```

Query Match 26.1%; Score 273.8; DB 10; Length 840;
Best Local Similarity 66.1%; Pred. No. 1.3e-57;
Matches 392; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

```

QY 300 AGGAATTTTCCAGCAATGAGGTTGGTGTAGATTTTGTATAGCAAGACTGCTACTGAG 359
    |||||
Db 727 ATACATTTGGCTGTATGGAAGTTAGTGTCACAAATTTCTTATGACGAGCTTTGATGCT 668
QY 360 GTTGATTAAGCGAGCAATGCAAGCTTATTAAGTTCTTGATTAACAGAGATGATCGGA 419
    |||||
Db 667 GTAGAGAAAGCTGCACAAAGCNCCTTACAAATCTCCAGAAAAGAGCAGCATGATG 608
QY 420 ATAGCTTTTGGCTGTGATATGATGATGAGCAAGTTTGAAGAGGTTTCTCCCG 479
    |||||
Db 607 CAACATGCAATTTGATTTGATGATGATGAGTGAACCCNMCNTGAGAAAGGTGTTCTCTCC 548
QY 480 GGGAGGTTGCGCATCTCCAGATATGATATGATATTAATTTGATGATTTATGATAT 539
    |||||
Db 547 GGAAGGTAGCACTATGACAGATATGTGTGACACATGATTTGATTTCTTACATCTTA 488
QY 540 TTTCAATTCGTATCCCTCAAGCTCCACACATTTATTAAGATCAACACTTGAAG 599
    |||||
Db 487 ATTCATTTGGAATCCCTCAAAATTTACAGCTTTCTTGAAGATCCACAGCTTGAAG 428
QY 600 GTTAGTATTGGAATTTATGCTGCTGTGAGAGCTTTTCCATGACATATGAGATGATATC 659
    |||||
Db 427 GTTGGAGCTGGGATGATGATGCTGTGAAGTTTATTAAGATTAATATATCTGTT 368

```

OY	660	AAAGATGTAGGACCTTTGACATTTAGCCAAACCAAAAATTTGGTGGACATAAAAAATCG	719
Db	367	AAAGGTGTGACGGATCTTTCTTTATGCTAAATCAAAAGCTTGGTGACATCATAAAGTGG	308
OY	720	GCCCTTGGCCTCACTAACGAGACACTTGTGTCGAAGAAGCTCCTAAGCAACAGAAATC	779
Db	307	GCTTGGCATCTTTGACGCAAAAACCTTCTATCAAAACAGCTTAAAAAGGCCAACAAATA	248
OY	780	AGGCTTGGGAACGTGGAGCTTTATCCTCTGTCAAAAGCAGCACTTCAATATGCGACCAACG	839
Db	247	AGACTGGGAAAATTGGGAGGCTCCTCTTTTGTCAAAAGGAGCACTAGAGATATCTGCCAACA	188
OY	840	GATGCTTATGCTCATCGATCCTTATACAAAGTTCCTTAAGGACCTTCGTATCG	892
Db	187	GATGGCTTTTGGCTTCTTGCTGTCTTATATACGGATTAAGATCTCCCGGACGC	135

RESULT 3	LOCUS	DEFINITION
BG441595	BG441595	704 bp mRNA linear EST 15-MAR-2001 GA_Ea0013022f Gossypium arboreum 7-10 dpa fiber library Gossypium arboreum cDNA clone GA_Ea0013022f, mRNA sequence.

ACCESSION	BG441595
VERSION	BG441595.1
KEYWORDS	GI:13351247
SOURCE	EST.
ORGANISM	Gossypium arboreum.
	Gossypium arboreum

REFERENCE
1 (bases 1 to 704)

AUTHORS Mang, R.A., Frisch, D., Yu, Y., Maim, D., Rambo, T.T., Simmons, J., Henry, D., Wood, T.C., Leslie, A. and Wilkins, T.A.

TITLE An integrated analysis of the genetics, development, and evolution of the cotton fiber

JOURNAL unpublished (2000)

COMMENT

Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293

source	Email: twing@clemons.edu Seq primer: TAATACGACGCTCAGTCATAGG High quality sequence stop: 701. Location/Qualifiers 1..704
--------	--

```

/organism="Gossypium arboreum"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ea0013022f"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/tissue_type="Fibers isolated from bolls harvested 7-10 dpa"

```

```

/lab_host="E. coli"
/note=Vector: pBR-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT      188 a      144 c      155 g      217 t
ORIGIN
Query Match      24.9%; Score 261; DB 12; Length 704;
Best Local Similarity 63.7%; Pred. No. 1.9e-54;
Matches 434; Conservative 0; Mismatches 23; Indels 12; Gaps 2;

```

QY	145	24	QY	205	84
CTTCTTCGTCGTCGACCGTAGCAAGGTACAACTCCGTCCTCATGGCCACGAGGAGATC	CTTCTTCGTCGTCGACCGTAGCAAGGTACAACTCCGTCCTCATGGCCACGAGGAGATC	CTTCTTCGTCGTCGACCGTAGCAAGGTACAACTCCGTCCTCATGGCCACGAGGAGATC	CAATCAATCAATCCCAATATATCTCGTCGCCAATTCCTCGTTCATCACTTCTTTAAT	ATCATCAAGATTAACCTCTCAAAACACACGCCCAATTCCTCATTTCTCTCATCCCGGC	
204	83	264	143		

OY	265	CTTATTAAGCATTTCTCTCTCCCTCCGCGACATA-----GGAATTTCCAGCAA	315
Db	144	CTTCTTTTCCGCTCTCTCCCTCGCTGTCATTAATCAAGCCTTAGATATCCGCCGT	203
OY	316	TGAGGTTGGTGCGTAGATTTTGTATPACCAAGACTGCTACTGAGGTGATTAAGCAGCAA	375
Db	204	TGAGATTTTGAGAGTCATATTTTGTATACCTTTACGGAAAGATGAGGTAGAAAAATGCTGCAA	263
OY	376	TGCAGCTTATTTAAGTCTTGTATACCAAGAGAGATGAATCTGGAAATPACCTTTTGTGGCT	435
Db	264	TGGACCTTATTAATAATTTGTGAATAATAGAGAAAGAAATGGGTCACGTGCTTTAGGGT	323
OY	436	TGGATATTGAGTGGGAGACCAAGTTTAAAAAAAGTGTCCTCCCGGGAAAGTTGGAGCATG	495
Db	324	TTGATATTGAGTGGAGAGCCCTCTTCTCAAAAAGAAATTTTGGCTGGGAAGCGTGCGTGA	383
OY	496	TTCAGATATTGCTGAATATGATATTAATTTGTGATGATATCATATTTTATTCCTGATATCC	555
Db	384	TGCATATATTGTTGTACACAGTCAGTATTGTGTATGTATATGCATATTTTTCATTCCTGCAATAC	443
OY	556	CTCAAAAGTCTCCACATCTTATTGAAGATTCAACCTTTGAAAGGTAGTATTGGAATTG	615
Db	444	CTCAAGTCTGAGAGTCTCTCTTGGAGCATCCGAATATTAAAAATTTGAGATTGCAATTG	503
OY	616	ATGTCGACTCTGTGAAGCTTTCCATGACTATGAGGTTAGTATCAAAATGTTGAGGATC	675
Db	504	ATGGGAGTGTCTCAAGGTGTTCACTGACTATAAAGATATCTGTTATATCTTTGGAAAGATC	563
OY	676	TTTGCATTTAGCCCAACCAAAAAATTTGGGAGAT---AAAAATGGGGGCGCTTGGCTCAC	732
Db	564	TTTTCGATCTTACCGCATTAATAATTTGGTAGTAATTTGCCCGGCACCTGAGATTTGCTGCTC	623
OY	733	TAACTAGACACATTTTGGCAAAAGACTCTGAAGCCCAACAGAAATCAAGCTTTGGGAAT	792
Db	624	TGACAGAGAGATTTATTGGCAAAAGAGCTTCCGAGACCCCAAGAAATTAAGCTGGAAACT	683
OY	793	GGGAGTTTATTCCTCTGTCAA	813
Db	684	GGGAGTTATATCTTTATCA	704

RESULT 4	AM201789	LOCUS	DEFINITION
AM201789	552 bp	mRNA	linear EST 02-DEC-2001
sif7c10.y1	Gm-c1027	glycine max	CDNA clone
Gm-c1027-1579	5'	similar to	TR:Q92242 Q92242 WRN PROTEIN. ; mRNA
sequence.			

ACCESSION	AW201789	
VERSION	AW201789.1	GI:6482532
KEYWORDS	EST.	
SOURCE	soybean.	
ORGANISM	Glycine max	

Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta, Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae: eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
I (pages 1 to 352)				
Shoemaker R.	Kalm, P., Vorkin, L., Erpelding, J., Correll, V., Khanna			
	A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,			
	Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers			
	, Y., Peterson, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk			
	, R., Ratter, E., Kohn, S., Shih, T., Jackson, Y., Cardenas, M., McCann			
	, R., Waterston, R. and Wilson, R.			
	Public Soybean EST Project			
Unpublished (1999)				
Contact: Shoemaker	R/Public Soybean EST Project			

Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel.: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available through: Reagen, Invitrogen Corp. 2130

Db 515 TTTCAGATTAGCCAAACAAAATTTGGTGGAGATTAATAATGGGCGCTTCCTCACTAA 574

QY 736 CTGAGACACTGTTGTTGCAAGAGCTCTGAGCCAAACAGATCAGGCTTGGGAAGTGGG 795
|||||
Db 575 CTGAGACACTGTTGTTGCAAGAGCTCTGAGCCAAACAGATCAGGCTTGGGAAGTGGG 634

QY 796 AGTTTATCTCTGTCAAGCAGATTACATACGC 832
|||
Db 635 AGGTCATNCTCTATCAAGCAGCAGTACATNCGC 671

RESULT 6
AM832139 467 bp mRNA linear EST 03-DEC-2001
LOCUS sm20b11.y1 Gm-c1027 glycine max cDNA clone GENOME SYSTEMS CLONE ID:
DEFINITION Gm-c1027-9430 5' similar to SW:WRN_HUMAN Q14191 WERNER SYNDROME
ACCESSION AM832139
VERSION AM832139
KEYWORDS EST.
SOURCE soybean.
ORGANISM glycine max
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots:
Rosidae: eustosids I: Fabales: Fabaceae: Papilionoideae: Phaseoleae:
glycine.
1 (bases 1 to 467)
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, D., Coryell, V., Khanna
A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers
Y., Person, B., Smaller, T., Gibbons, M., Pape, D., Harvey, N., Schuk
R., Rittler, E., Kohn, S., Shiu, T., Jackson, Y., Cardenas, M., Mccann
R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccuteresgen.com
Insert Length: 734 Std Error: 0.00
High quality sequence stop: 417.
Location/Qualifiers
1..467
/organism="glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1027-9430"
/clone_lib="Gm-c1027"
/tissue_type="cotyledons of 3- and 7-day-old Williams
seedlings"
/lab_host="DH10B"
/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
XhoI; This cDNA library was constructed from mRNA isolated
from cotyledons of 3- and 7-day-old Williams seedlings
which were propagated on paper towels with distilled
water. The cotyledons were flash-frozen in liquid
nitrogen, then lyophilized for 72 hours. Unequal amounts
of mRNA was used for cDNA synthesis. Stratagene's cDNA
Synthesis Kit (catalog number 200401) was used to
synthesize the cDNA. First-stranded synthesis was
performed with 5-methyl dCTP, hence the ligated cDNA was
hemimethylated. A modification of Stratagene's
first-strand synthesis primer was used. An anchor
nucleotide (V=A, C, or G) was added to the 3' end of the
primer [GAGAGAGAGAGAGAGAGAGACTACTGCGAG(T)18] to anchor
the primer at the 5' end of the poly(A) tract. After
second-strand synthesis, the cDNA ends were filled in
with cloned Pfu DNA, ligated to EcoRI adapters and

subsequently phosphorylated. The XhoI site within the
first-strand synthesis primer was then restricted by
digestion with XhoI; all XhoI sites in the cDNA would be
protected by their hemimethylated status. The cDNA
constructs were size-fractionated with a 500 bp cutoff,
using GibcoBRL Life Technologies' cDNA Size Fractionation
column. The column eluent was then ligated into
Stratagene's pBluescript(tm) II XR predigested vector
(pBluescript II SK(+)) that has been digested with EcoRI
and XhoI, and phosphorylated by Stratagene). 97% of the
white and blue colonies appear to contain recombinant
plasmids with cDNA inserts, based on size (n=30). This
library was constructed by Dr. Paul Keim and Dr. Virginia
Coryell."

BASE COUNT 133 a 84 c 109 g 140 t 1 others
ORIGIN

Query Match 23.0%; Score 241.2; DB 10; Length 467;
Best Local Similarity 70.0%; Pred. No. 1,6e-49;
Matches 324; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

QY 430 TTGGCTTGATATTGAGTGAGAGACCAAGTTTACAAAAGGTGTCGCCGGGAAGGTTG 489
|||||
Db 4 TTGGATTGTGACATTGATGAGAAACCCACCTTCAGAAAAGGTGTCGCCGGGAAGGTTAG 63
|||||

QY 490 CGACTGCCAGATATGTGTAGATAGTAATTATTTGATGATTATTTTCATTCTG 549
|-|
Db 64 CAGTGATGACATATGTGTGACACTAGACATTTGTCATGTTCTACATCAATTCATTCTG 123
|||||

QY 550 GTATCCCTCAAGACGTCCACATCTTATGGAAGTTCAACACTGTAAAGTAGTATG 609
|||||
Db 124 GAATCCCTCAAAATTTACAGCTTTTGGCTTGAAGATCCACAGCTTTAAAGGTGGAGCT 183
|||||

QY 610 GAATTGATGTGACTGTGCAAGCTTTTCCATGCTATGAGTAGTATCAAAAGATGTTG 669
|||||

Db 184 GGATTGATGTGATGATGCTGTGAAGGTTTTAGAGTTATTAATCATATCTGTTAAAGTGTGA 243
|||||

QY 670 AGGATCTTTACAGATTAGCCAAACAAAATTTGGTGGAGATTAATAATGGGCGCTTCCT 729
|||||

Db 244 CGGATCTTTCTTTTCATGCTATCAAAAGCTTGGTGAGATCAATAGTGCGGCTTCAT 303
|||||

QY 730 CACTTAAGTACACTGCTGTTTGGCAAGAGCTCCGAGCCCAACAGATACAGGTTGGGA 789
|-|

Db 304 CTTGACTGAAAACCTCTATCAACAGCTTAAAGGCCCAACAAATTAAGACTGGAA 363
|||||

QY 790 ACTGGAGTTTATCTCTGTCAAGCAGATTACATACGACGACGATGCTATG 849
|-|

Db 364 ATTGGAGGCTCCGTGTTTGTCAAGGACGACACTAGATGCTGCAACAGATGCTTTTG 423
|||||

QY 850 CTTGATGGCATCTTTACAGGTTCTTAAGGACCTTCTGATGC 892
|||||

Db 424 CTTCTTGTTGTCTTATCAGGCGATTAAGATCTCCGGAGGC 466
|||||

RESULT 7
B0799068 661 bp mRNA linear EST 30-JUL-2002
LOCUS B0799068
DEFINITION EST 1237 Green Grape berries lambda zap II library Vitis vinifera
ACCESSION B0799068
VERSION B0799068.1 GI:22014034
KEYWORDS EST.
SOURCE Vitis vinifera.
ORGANISM Vitis vinifera.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots:
Vitaceae: Vitis.
1 (bases 1 to 661)
Abbal, P., Agasse, A., Ageorges, A., Atanassova, R., Barrieu, F.,
Couture, C., Dedalechamp, F., Delrot, S., Glissant, D., Grimplet, J.,
Hamdi, S., Romieu, C. and Terrier, N.
Generation of Expressed Sequence Tag from Grape Berry (skin, pulp
or seeds) at Various Developmental Stages

RESULT 9
LOCUS BG133220 506 bp mRNA linear EST 31-JAN-2001
DEFINITION EST66112 tomato crown gall Lycopersicon esculentum cDNA clone
CT081B12 5' sequence, mRNA sequence.
ACCESSION BG133220
VERSION BG133220.1 GI:12633408
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; eunasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 506)
van der Hoeven,R., Sun,H., Cho,J., Uutterback,T., Hansen,C., Romning
C. and Tanksley,S. from tomato crown gall tissue
Generation of ESTs from tomato crown gall tissue
Unpublished (2001)
JOURNAL Contact: CGI
COMMENT Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>.
FEATURES
source
1. 506
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CT081B12"
/clone_lib="tomato crown gall"
/tissue_type="crown gall"
/dev_stage="crown galls from full-grown plants (8 wks old)"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:
XhoI; Four wk old greenhouse plants were stab inoculated
on stem with Agrobacterium tumefaciens C58 (Dr. T.J. Burr,
Cornell U.). Galls were allowed to develop for another 4
wks, when gall tissue was frozen in liquid nitrogen."
BASE COUNT 142 a 91 c 126 g 147 t
ORIGIN
Query Match 21.3%; Score 223; DB 12; Length 506;
Best Local Similarity 65.2%; Pred. No. 5.6e-45;
Matches 328; Conservative 0; Mismatches 175; Indels 0; Gaps 0;
QY 363 GATACGAGCAATGACGCTTATTAAGTTCTTGATACCAAGAGATGATCGAATA 422
DB 2 GAGAAATCTGCACTAGAGCTGTTAAATTTGTTGAGGAAAGAAAGAAAGCAAT 61
QY 423 GCTTTGTTGGCTGGATATATGATGAGAGACCAAGTTTGAAGAGGTCTCCCGGG 482
DB 62 GTTGCTTTGGATTTGACATTTGATGAGAGCCACCTTTTGAAGAGGTGCGACCTGG 121
QY 483 AAGTTGCGACTGTCAGATATGTAGATAGTATATTTGATGATGATGATATTTT 542
DB 122 AAGCGTGTGTTATGACATATGTGTGACAAAGGTAATTTGATTTGATATCATC 181
QY 543 CATTTGATGATCCCTCAAGTCTCCAAACATCTTATTTGAAGATTCACACTTGA 602
DB 182 CACTCGGAGATCCCTCAACTCTGCAATCTCTTCTGAGATTCACACTGTTGGAAGTG 241
QY 603 GGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 662
DB 242 GGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 301
QY 663 GATGTTGAGATCTTACAGATTAGCCAAACAAATAATTTGGTGAAGATTAATAA 722
DB 302 GCTTTGGAAGATCTTCCGCACTTGCCAAACAAATAATTTGGTGAAGATTAATAA 361

QY 723 CTTCCTCACTAAGTACGACACTTGTTCGAAGAGCTCTGAGACCAACAGATCAG 782
DB 362 CTAGCATCATTAAGTACGAGAGCTTCTTGCCAGACAGCTCCCAAGCAAGTATATCAG 421
QY 783 CTGGGAGACGAGGAGTTTATCTCTGTCGAAGACAGAGTTACAAATCCAGACAGAT 842
DB 422 TTGGGGAATTTGGAGGCTTATGTTATCTAGGACCACTACATTAATGCTGATCAGAT 481
QY 843 GCTTATGCTTCATGAGCATCTTTA 865
DB 482 GCCTTTGTTCCGTGACTATA 504
RESULT 10
LOCUS BM094501 531 bp mRNA linear EST 30-NOV-2001
DEFINITION sa16q03.y1 Gm-cl066 Glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-cl066-3005 5' similar to TR:Q9SYM6 Q9SYM6 HYPOTHETICAL 35.1
KD PROTEIN.; mRNA sequence.
ACCESSION BM094501 GI:17023467
VERSION BM094501.1
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE 1 (bases 1 to 531)
Shoemaker,R., Keim,P., Vodkin,L., Erpelidng,J., Corryell,V., Khana
A., Bolla,B., Matra,M., Hillier,L., Kuebda,T., Matlin,J., Beck,C.,
Wyle,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk
R., Riter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
R., Waterston,R. and Wilson,R.
Public Soybean EST project
Unpublished (1999)
JOURNAL Contact: Shoemaker R./Public Soybean EST Project
COMMENT Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntville, AL 35801 For further information
call: (800)-533-4363 or contact via email: coueresgen.com
High quality sequence stop: 423.
FEATURES
source
1. 531
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl066-3005"
/clone_lib="Gm-cl066"
/tissue_type="leaf and shoot tip, salt stressed, 2 week
old seedling"
/lab_host="DH10B"
/note="Vector: pBluescript II SK+; Site.1: EcoRI; Site.2:
XhoI; The cDNA library was constructed from mRNA isolated
from unexpanded leaves and the shoot tips of 2 week old
seedlings from the cultivar Williams. The 2 week old
seedlings were salt stressed in a solution of 500mM NaCl
for 3 days prior to harvesting. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with a XhoI restriction site. EcoRI
adapters were ligated to the blunt-ended cDNA fragments
followed by XhoI digestion. The cDNA fragments were
directionally cloned into the EcoRI-XhoI restriction site
of the pBluescript vector. The ligated cDNA fragments were
transformed into DH10B host cells (GibcoBRL). This library
was constructed in the laboratory of Dr. Randy
Shoemaker."

ORIGIN

Query Match 20.4%; Score 213.6; DB 13; Length 531;
Best Local Similarity 64.6%; Pred. No. 1.2e-42;
Matches 340; Conservative 0; Mismatches 174; Indels 12; Gaps 1;

QY 299 TAGAATTTCCAGCATAGTTGGTGTAGAGATTTTGTATAGCAAGACCTGCTACTGA 358
16 TATGACATTTCCCTCTATATAGAGTTTAGTGTCAAAATTTTATATACAGACTTTTATGATC 77
QY 355 GGTGATTAACGAGCAATGACGCTTATTAAGTTCTTGATACAGAGAGATGAATCTGG 418
78 TGTAGAGAAAGCTGCAACAAGCTTACAAATTCCTCAAGAGATGATGCA----- 128

QY 419 AATAGCTTTTGTGGCTGTGATATTTAGTGGAGACAGTTTATAGAAAGCTTCTCC 478
129 ---AATTGCAATTTGATTTGATGATGAGAAACCACTTGAGAAAGGTGTTCAC 185

QY 479 GGGGAAGTTGCCAGCTCTCCAGATATGTGATAGTAAATTTATTTGATGATAT 538
186 CGGAAGAGTACAGATGATGAGATATATGTGACACATGATTCATCTTCTACATCT 245

QY 539 TTTTCAATCTGTATCCCTCAAGTCTCCACATCTTATTTGAAGATTCAACACTGTAA 598
246 AATTCAATCTGTGAATCCCTTGAATTTACAGCTTTTGTGAAGATCCACAGCTTTGAA 305

QY 599 GGTAGGATTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 658
306 GGTGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 365

QY 659 CAAAGATTTGAGATCTTTCAGATTTAGCCAAACAAATTTGTGAGATTAATAAATG 718
366 TAAAGGTGACGAGATCTTCTTTTCATGCTAATCGAAAGCTTGTGAGATATAGTG 425

QY 719 GGGCTTTGCTCTACTACTAGACACTGTGTGCAAGAGCTCTGACCAACAGAT 778
426 GGTCTTCTGATCTTGTACTGAAACCTTTTATGCAAAACCTTAAAGCTTAAACAT 485

QY 779 CAGCTTGGGAGCTGGAGTTTATCTCTGTCAAGACAGCAGTTA 824
486 AAGACTGGGAATTTGGAGACTCTCTGTTTGTCAAGAGCAACTA 531

RESULT 11
BM358892 515 bp mRNA linear EST 09-JAN-2002
LOCUS GA_Ea0013P02r Gossypium arboreum 7-10 dpa fiber library Gossypium
DEFINITION
arboeum cDNA clone GA_Ea0013P02r, mRNA sequence.
ACCESSION BM358892
VERSION BM358892.1 GI:18099638
KEYWORDS EST.
SOURCE Gossypium arboreum.
ORGANISM Gossypium arboreum.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosids II; Malvales; Malvaceae; Gossypium.
1 (bases 1 to 515)
Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
D., Wood,T.C., Leslie,A. and Wilkins,T.A.
An integrated analysis of the genetics, development, and evolution
of the cotton fiber
Unpublished (2000)
CONTACT: Wing RA
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: twing@clemson.edu
Total High Quality bases = 415
Seq primer: TAATAGCACTCACTATAGGG
High quality sequence stop: 514.
Location/Qualifiers

source 1.. 515
/organism="Gossypium arboreum"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ea0013P02r"
/clone_1b="Gossypium arboreum 7-10 dpa fiber library"
/tissue_type="Fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/note="Vector: pBR-CMV; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 134 a 112 c 109 g 160 t

ORIGIN

Query Match 16.4%; Score 172.2; DB 13; Length 515;
Best Local Similarity 64.6%; Pred. No. 2.6e-32;
Matches 277; Conservative 0; Mismatches 143; Indels 9; Gaps 1;

QY 208 ATCAATCCCAATTAATCCGTCGCAATTTGCTTCATCATCTCTTCTACATCTT 267
87 ATCAAGATTAACCTCAAAACAGCGCCCAATTTGCCATTTCTATCTCACCCGCCCTT 146

QY 268 ATTAACGATTTCTCTCTCCGTTGCCAGCTA-----GGAATTTTCCAGCAATGA 318
147 CTTTCCGCTCTCTCTCCGCTGCTCAATCTAATCAAGCTTATGATATCCGCCCTTGA 206

QY 319 GGTGCTGTGATGATTTGTATAGCAAGCTCTGATGATGATGATGATGATGATGAT 378
207 GATTTGGAGCTATATTTGTATAGCTTTTACGAGAGATGAGGAGAAATGCTGCAATGG 266

QY 379 AGCTTTTAATAGTTCTTGTATACCAAGAGATGAATCTGAAATGCTTTTGTGCTTGG 438
267 AGCTATTAATAATTTGTGAATTAAGAAAGAAATGAGTCAATGCTTTAGGTTTG 326

QY 439 ATATTAGTGGAGACCAAGTTTGAAGAGGTCTTCCCGGGAAGTTGGACATGTC 498
327 ATATTAGTGGAGAGCCCTCTTCCAAAGAAATTTTGTGCTGGAGGCTGGGCTGATGC 386

QY 499 AGATATGTAGATAGTAAATTTATTTGATGATTTATGATATTTTTCATCTGATCCCTC 558
387 AGATATGTGTGACATGATGATTTTATTTATTTATTTTTCATCTGATCCCTC 446

QY 559 AAGTCTCCACATCTTATTTAGATTTCAACACTTTTAAAGGTAGTATTTGAATTTGATG 618
447 AAGTCTGACAGTTCTTGTAGAGACTCGAAATTTATTAAGTTTGAATTTGATTTGATG 506

QY 619 GTGACTCTG 627
507 GCGATGCTG 515

RESULT 12
BM358891 517 bp mRNA linear EST 09-JAN-2002
LOCUS GA_Ea0013022r Gossypium arboreum 7-10 dpa fiber library Gossypium
DEFINITION
arboeum cDNA clone GA_Ea0013022r, mRNA sequence.
ACCESSION BM358891
VERSION BM358891.1 GI:18099637
KEYWORDS EST.
SOURCE Gossypium arboreum.
ORGANISM Gossypium arboreum.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosids II; Malvales; Malvaceae; Gossypium.
1 (bases 1 to 517)
Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
D., Wood,T.C., Leslie,A. and Wilkins,T.A.
An integrated analysis of the genetics, development, and evolution
of the cotton fiber
Unpublished (2000)
CONTACT: Wing RA
Clemson University Genomics Institute

[illegible]

RESULT 14	
Bj213300	
LOCUS	Bj213300 595 bp mRNA linear EST 04-APR-2002
DEFINITION	Bj213300 Y. Oglthara unpublished cDNA library, Wh Triticum aestivum
ACCESSION	CDNA clone wh21m08 5', mRNA sequence.
VERSION	Bj213300
KEYWORDS	Bj213300.1 GI:19952807
SOURCE	EST.
ORGANISM	bread wheat.
	Triticum aestivum

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 595)	Ogihara,Y and Murai,K	Expressed genes in <i>Triticum aestivum</i>	Unpublished (2002)	Contact: Tadasu Shin-1

Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.

FEATURES	Location/Qualifiers
source	1. .595

```

/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="vh2lm08"
/clone_1lb="Y. Ogihara unpublished cDNA library, Wh"
/lssue_type="spike at meiosis"
/dev_stage="Feekes' scale 9"
/notes="Vector: Lambda Uni-ZAP XR, excised phagemid:
Site_1: EcoRI, Site_2: XhoI. Plants were grown under
hydroponic conditions at UC Davis, salt stressed for 12
hours, and for 7 days, then dissected and frozen (Akunov
in J Dvorak Lab). Total RNA was prepared from sheath
tissue, equal quantities of RNA were pooled from the two
samples, polyA was purified from the pooled RNA, a cDNA
library was made, and the cDNA clones were in vivo
excised to give pluescript phagemids in the TV Close lab
at the University of California, Riverside (Akunov, Chin
, Choi, Close, Fenton, Kianian, Otto, Simons, Zhang).
Plasmid DNA preparations and DNA sequencing were
performed in the OD Anderson lab (all other authors)."

```

Query Match	15.4%	Score 162	DB 13	Length 595
Best Local Similarity	58.3%	Pred. NO. 9,1e-30		
Matches 323	Conservative	0	Mismatches 225	Indels 6
				Gaps 2

[illegible]

Db	105	ATTGACGCAAAATGGAAGCATGA - --AGGCGCTTGCCCGCGCTCCCTTGATTTGCAT	161
Qy	441	ATTGAGTGGAGACCAAGTTTAGAAAAGTGTCTCCCGGGGAAAGTTGCCATGTCCAG	500
Db	162	CTCGGTTGGAAACCCCTTTTCCAGAGAGGAACCCACCATGTAAAGTCGCTGTAGTCGA	221
Qy	501	ATATGTGATAGATAGTAATATTGTGATGTTATGCAATATTTTTCATCTTCGATGCCTCGA	560
Db	222	TTATGCAATGGAACAAACATCATTTTATCTCATGATATCATTCATCTCGGGGTGCCTCC	281
Qy	561	AGCTCCCAACATCTTATTGGAAGATTCAACACTTGTAAAGTAGTATTGGAAATGATGCT	620
Db	282	ATCTTGAATATCTCTTTTGAGAGACAGTTCAACCGTTAAAGTTGAGATGTATAGACAA	341
Qy	621	GACTCTGTGAAGCTTTTCCATGACTAGTAGAGTATGATCAAGAATGTTGAGATCTTTCA	680
Db	342	GATCGAAGGAAATGTTCAATGATTAATGATATGTCCTGTACAAACCATGATGATTTATCA	401
Qy	681	GATTTAGGCAACCAAAAAATTTGGTGG - -AGATAAAAAATGGGGCCCTTGCTCACTA	737
Db	402	ACTGTGGCAACGTCMAATGATAGCTGGGCCCTATAAAGATGAGATCTTGCTGCAATTA	461
Qy	738	GAGACACTGTTTGGCAAGAGCTCCTGTGAAGCAACAGATCAAGCTTGGGAACCTGGAG	797
Db	462	GAAATGTTACATGTAAAGAGTTGCCAAGGCTGGCAACATTAAGAAATGGAAACTGGAG	521
Qy	798	TTTTATCCTCTGTCAAGACAGCACTTACAAATACGACGCAACGAGTCTTATGCTTCA	857
Db	522	TCTTTTGTCTCTCAAAAAACCACTTGATGATGCTGTACCGATGCTTACATCTCTGG	581
Qy	858	CATCTTTCAAGGT 871	
Db	582	TACTTTGATAGAGT 595	

RESULT 15	BI433038	LOCUS	DEFINITION
	BI433038	499 bp	mRNA
	EST533759	P. infestans-challenged leaf	linear EST 21-AUG-2001
	PP4286	'5' sequence, mRNA	Solanum tuberosum cDNA clone

ACCESSION	BI433038	:
VERSION	BI433038.1	GI:15257728
KEYWORDS	EST.	:
SOURCE	potato.	:

ORGANISM	REFERENCE
<i>Solanum tuberosum</i>	1 (bases 1 to 499)
Eukaryotes: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:	
Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots:	
Asteridae: euasterids I: Solanales: Solanaceae: Solanum.	

AUTHORS Kestrepo, S., Grillitsis, H.M., Smart, C.D., Cho, J., Cienmungo, A., Bougiti, O., Buell, C.R., Roming, C.M., Fry, W.E. and Baker, B.

TITLE Generation of ESTs from Potato leaves Challenged with *Phytophthora infestans*, Compatible Interaction

JOURNAL unpublished (2000)

CONTACT: Cathy Rinning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: M13P-R
book@cdna.resgen.com

FEATURES	location/Qualifiers
source	1..499

```

/organism="Solanium tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="PPCA286"
/clone_lib="p_infestans-challenged_leaf"
/tissue_type="leaf"
/dev_stage="6 week old"
/lab_host="SOLR"
/node_vector: bluescript SK(-): Site_1: Ecori: Site_2:
XhoI: Whole plants were challenged with 450,000
sporangia/ml P. infestans US-1(US 940501) in Biotron
(Madison, Wisconsin). Leaf tissue was collected at 1, 2,

```

5, 12, and 24 hours post-challenge and frozen in liquid nitrogen immediately upon removal. Kennebec plants showed no signs of HR. Katahdin plants (susceptible to P. infestans US-1) were used as controls and showed infection. NOTE: We cannot exclude the possibility that this sequence is actually derived from Phytophthora rather than potato."

BASE COUNT 141 a 86 c 118 g 154 t
ORIGIN

Query Match 14.9%; Score 156.8; DB 13; Length 499;
Best Local Similarity 62.5%; Pred No. 1.8e-28;
Matches 245; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

```
QY 296 ACCTAGGAATTTTCCAGCATGAGTTGGTGTAGATTTGTATAGCAAGACTGCTAC 355
    || || || || || || || || || || || || || || || || || || || ||
Db 108 AGTATGCAATACCCGGAATGCTTTTAAAGACATGTTATATACAGACGAACTATTAA 167

QY 356 TGAGGTTGATTAAGCGAGCATGCACTTATTAAGTTCTTGATACCAAGAGATGATC 415
    || || || || || || || || || || || || || || || || || || || ||
Db 168 AGAGGTGGAGAAATCTGCAGTAGAGCTGTTAAATTTGTTAGGAAAGAGAAAGA 227

QY 416 TGGATAGCTTTTGTGGCTTGATATGAGTGAGACCAAGTTTAGAAAGGTGTC 475
    || || || || || || || || || || || || || || || || || || || ||
Db 228 AGGGAATGTTGCTCTTGATTTGACATGATGAGAGCCOACTTTAGAGAGGTGCGC 287

QY 476 CCCGGGGAAGGTTGCGACTGTCAGATATGTAGATAGTAATATTATGATGATGCA 535
    || || || || || || || || || || || || || || || || || || || ||
Db 288 ACCTGGGAGGCGCGTGTATGAGATATGTGTGACAAAGGCTAATTGTTATCTGCA 347

QY 536 TATTTTCATTTGCGTATCCCTCAAGTCTCCACATCTTATTGAAGATTCAACTGT 595
    || || || || || || || || || || || || || || || || || || || ||
Db 348 TATCATCCACTCTGGAAATCCCTCAAACTCTGCAATCTCTTGTGAGATCCAACTGTGT 407

QY 596 AAAGTAGGTATTTGAATTGATGTGACTCTGGAAGCTTTTCCATGATGATGAGTTAG 655
    || || || || || || || || || || || || || || || || || || || ||
Db 408 GAAGTGGGTGTTGATGATGCAAAATGATGCTTACAAAGTTCCGCAAGATCACAATGATC 467

QY 656 TATCAAAAGATGTTGAGGATCTTTCAGATTAG 687
    || || || || || || || || || || || || || || || || || || || ||
Db 468 TGTGAAGCTTTGGAAGATCTTCTGAACCTTG 499
```

Search completed: November 2, 2002, 15:31:41
Job time : 2052 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 2, 2002, 15:06:14 ; Search time 52 Seconds

(without alignments)
6720,758 Million cell updates/sec

Title: US-09-896-186b-23

Perfect score: 1049

Sequence: 1 accaagcattatcttattc.....tttaacgcgcagaactag 1049

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 310279 seqs, 166577418 residues

Total number of hits satisfying chosen parameters: 620558

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	111	10.6	5189	10 US-09-954-456-1131	Sequence 1131, App
2	44.4	4.2	3149	10 US-09-729-674-127	Sequence 127, App
3	38.2	3.6	439	10 US-09-864-761-20174	Sequence 20174, A
4	36	3.4	99916	10 US-09-816-095-3	Sequence 3, Appli
5	36	3.4	659158	9 US-09-771-208-20	Sequence 20, Appli
6	35.8	3.4	29544	10 US-09-464-767-1	Sequence 1, Appli
7	35.8	3.4	32745	10 US-09-464-767-3	Sequence 3, Appli
8	35	3.3	954	10 US-09-947-971-3	Sequence 3, Appli
9	34.6	3.3	280	9 US-09-933-797-413	Sequence 413, App
10	34.4	3.3	463	10 US-09-864-761-5350	Sequence 5350, App
11	34.4	3.3	53326	10 US-09-818-264-3	Sequence 3, Appli
12	34	3.2	527	10 US-09-864-761-6599	Sequence 6599, App
13	34	3.2	527	10 US-09-864-761-8066	Sequence 8066, App
14	34	3.2	1827	10 US-09-925-302-270	Sequence 270, App
15	33.8	3.2	556	10 US-09-864-761-8619	Sequence 8619, App
16	33.8	3.2	785	10 US-09-770-445-860	Sequence 860, App
17	33.6	3.2	258	10 US-09-878-574-10134	Sequence 10134, A
18	33.6	3.2	285	10 US-09-878-574-7317	Sequence 7317, App
19	33.6	3.2	5814	10 US-09-764-847-1860	Sequence 1860, App

C	20	33.4	3.2	533	10	US-09-864-761-6759	Sequence 6759, App
	21	33.4	3.2	2979	10	US-09-815-242-4754	Sequence 4754, App
	22	33.4	3.2	3009	10	US-09-815-242-8801	Sequence 8801, App
	23	33.4	3.2	3009	10	US-09-815-242-9060	Sequence 9060, App
C	24	33.4	3.2	7035	10	US-09-815-242-8615	Sequence 8615, App
	25	33.2	3.2	477	10	US-09-864-761-5436	Sequence 5436, App
	26	33	3.1	305	10	US-09-864-761-18262	Sequence 18262, A
	27	33	3.1	496	10	US-09-864-761-2534	Sequence 2534, App
C	28	33	3.1	1671	12	US-10-052-586-505	Sequence 505, App
	29	32.8	3.1	3294	10	US-09-764-860-1156	Sequence 1156, App
C	30	32.8	3.1	203654	10	US-09-820-905-3	Sequence 3, Appli
	31	32.4	3.1	471	10	US-09-795-668-1424	Sequence 1424, App
	32	32.4	3.1	471	10	US-09-795-668-1424	Sequence 1424, App
	33	32.4	3.1	1038	10	US-09-815-242-5161	Sequence 9161, App
	34	32.4	3.1	6109	10	US-09-880-107-3768	Sequence 3768, App
C	35	32.4	3.1	198285	10	US-09-880-107-3814	Sequence 3814, App
	36	32.4	3.1	640681	10	US-09-790-988-1	Sequence 1, Appli
	37	32.2	3.1	1956	10	US-09-351-794A-1	Sequence 1, Appli
C	38	32.2	3.1	5340	10	US-09-817-360-1	Sequence 1, Appli
	39	32.2	3.1	21860	10	US-09-764-877-3803	Sequence 3803, App
C	40	32	3.1	153	10	US-09-924-035A-86	Sequence 86, Appli
	41	32	3.1	575	10	US-09-864-761-8070	Sequence 8070, App
C	42	32	3.1	1185	10	US-09-738-363-9	Sequence 9, Appli
	43	32	3.1	1185	10	US-09-977-653-9	Sequence 9, Appli
	44	32	3.1	1237	10	US-09-070-927A-555	Sequence 555, App
C	45	32	3.1	1577	10	US-09-864-761-30755	Sequence 30755, A

ALIGNMENTS

RESULT 1
US-09-954-456-1131
Sequence 1131, Application US/09954456
Patent No. US20020115057A1
GENERAL INFORMATION:
APPLICANT: Young, Paul
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
FILE REFERENCE: 689290-76
CURRENT APPLICATION NUMBER: US/09/954,456
CURRENT FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: US/60/233,617
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,923
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/235,637
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,134
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,637
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,638
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,711
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,720
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,840
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,863
PRIOR FILING DATE: 2000-09-27
NUMBER OF SEQ ID NOS: 2276
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1131
LENGTH: 5189
TYPE: DNA
ORGANISM: Homo sapiens
US-09-954-456-1131
Query Match 10.6%; Score 111; DB 10; Length 5189;
Best Local Similarity 54.1%; Pred. No. 2.4e-22;
Matches 272; Conservative 0; Mismatches 225; Indels 6; Gaps 2;

OY	336	GATACCAGAGAGATGATCTGCGAATAGCTTTTGTGGCTTGATATGGAGACCA	4555
Db	433	GATATTACCATGAGCTCTATCAAGATGGGATGTGGTGGATTTGCACATGGAGTGGCACCA	4222
OY	456	AGTTTATAGAAAAGTGTTCTCCGGGGGAAGGTGCGACTGCCAGATATGTAGATAGT	5151
Db	493	TTATACAAATAGAGGGAACCT---TGCGAAAGTTGCACATTAATCAGTTGTGTGTTGAG	5459
OY	516	AATTATGTGATGTATTCATATTTTT---CATCTCGTATCCCTCAAGTCTCCACAT	5723
Db	550	AGCAAAATTTTACTTGTTCCAGCTTCTTCCATGTCACTGATTTTTCCCGCAGGATTTAAAAATG	6090
OY	573	CTTATTTAGATTCACACCTGTGAAAGTAGTATTTGGAATGTATGCTGCTGTGAAG	6322
Db	610	TTGCTTTGAAATTAACGACGTTAAAAAGGCAGCTGAGAAATTTGAAGAGATTCAGTGAAA	6659
OY	633	CTTTTCCATGACTATGGAGTTAGTATCAAGATTTGAGATCTTTTACATTTAGCCAAC	6522
Db	670	CTTCTACGTGACTTTGATATCAAAATTTGAAAGATTTTGTGGATGTACAGATGTTGCCAAT	7299
OY	693	CAAAAATTTGGTGGAGATAAAAAATGGGGCCCTTGCCCTCACAACTGAGACACTGTGTTGC	7522
Db	730	AAAAAGCTGAATGTATACAGNAGACCTGGAGCCTTAAACAGTCTGTTTAAACACCTTTAGGT	7859
OY	753	AAAGAGCTCTGAAAGCCACAACAGATCAGGCTTGGGAACCTGGGAGTTTATCCTCTGTCA	8122
Db	790	AAACAGCTCTGAAAGACCAACAGTCTATCCGCTGTGACATTTGGAGATTAATTTCCCTCCTACT	8494
OY	813	AAGCAGAGTTACAAATACGACAGCAAGGAGTCTTATCTTCATGCGACTTTTACAAAGTT	8722
Db	850	GAGGACGAGAAAGCTGATACGACGCCACAGTATGCTTATCTGTTTATTTATTTACCGAAAT	9099
OY	873	CTTAAGGACCTTCCGTGATGCTGT	895
Db	910	TTAGAGATTTTGGATGATACGT	932

RESULT 2
 US-09-729-674-127
 Sequence 127, Application US/09729674
 Patent No. US2001003935A1
 GENERAL INFORMATION:
 APPLICANT: Jacobs, Kenneth
 APPLICANT: McCoy, John M.
 APPLICANT: Lavallee, Edward R.
 APPLICANT: Collins-Racie, Lisa A.
 APPLICANT: Evans, Cheryl
 APPLICANT: Werberg, David
 APPLICANT: Treacy, Maurice
 APPLICANT: Agostino, Michael J.
 APPLICANT: Steinger II, Robert J.
 APPLICANT: Spaulding, Vikki
 APPLICANT: Wong, Gordon G.
 APPLICANT: Clark, Hilary
 APPLICANT: Fechtel, Hilary
 APPLICANT: Genetics Institute, Inc.
 TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
 FILE REFERENCE: 6055-64X
 CURRENT APPLICATION NUMBER: US/09/729,674
 CURRENT FILING DATE: 2000-12-04
 PRIOR APPLICATION NUMBER: 09/539,330
 PRIOR FILING DATE: 2000-03-30
 NUMBER OF SEQ ID NOS: 283
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 127
 LENGTH: 3149
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-729-674-127
 Query Match 4.28; Score 44.4; DB 10; Length 3149;
 Best Local Similarity 49.3%; Pred. No. 0.0069;

Matches	148:	Conservative	0:	Mismatches	146:	Indels	6:	Gaps	1:
OY	571	ATCTATTGAACATTTGCACACTCTTGAACCTTAGTATGTAATGTGCTACTCTGTGA							630
Db	470	AAATTTTGGCAGATGGCACCCATTTTGGAAAGTTGGAGTGGGATGCTCAGAAAGATGCCAGCA							529
OY	631	AGCTTTTCCATGACATATGGAGTTAGTATACAAAGATGTTGAGGATCTTTCAGATTAGCCA							690
Db	530	AGCTTCTGAGGATTTATGGCTCGTTGTATGAGGGGGCTCGCTGACCTCCATACCTAGCCA							589
OY	691	-----ACCAAAAAATTTGGTGGAGATATAAAAAATGGGGCCCTTCCTCATTACATCTGAGACAC							744
Db	590	TGCGGCAGAGAAACATTTGGCTCTGTATATGGGCTTAGCCTGAAGTCCCTCGCTGAGACTG							649
OY	745	TTGTTTGGCAAGAGCTCCTGAAGCCAAACAGAAATCAGGCTTGGGAACTGGGAGTTTATC							804
Db	650	TTTTTAACCTTTCCTTGGACAGTCCCTCTCTACTTCGTTGGACGAACCTGGATGCTGAAA							709
OY	805	CTCTGTCAAGCAGAGATACAAATPACGACGACAGCAAGCACTTATATGCTATATGAGCATCTTT							864
Db	710	CTCTCACAAGAGCAAGATTAATTTATGCTGCCAGGATATCCCAAGATTTTAGTGGCTCTCT							769

RESULT 3
 US-09-864-761-20174/c
 Sequence 20174, Application US/09864761
 Patent No. US20020048763A1
 GENERAL INFORMATION:
 APPLICANT: Penn, Sharron G.
 APPLICANT: Rank, David R.
 APPLICANT: Hanzel, David K.
 APPLICANT: Chen, Wensheng
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
 FILE REFERENCE: Aeomica X-1
 CURRENT APPLICATION NUMBER: US/09/864,761
 CURRENT FILING DATE: 2001-05-23
 PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: US 09/632,366
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: GB 24263.6
 PRIOR FILING DATE: 2000-10-04
 PRIOR APPLICATION NUMBER: US 60/236,359
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: PCT/US01/00666
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00667
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00664
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00669
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00665
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00668
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00661
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00670
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/774,203
 PRIOR FILING DATE: 2001-01-29


```
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1
SEQ ID NO 20174
LENGTH: 439
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL035419.9
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN LONG, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
US-09-864-761-20174

Query Match
Best Local Similarity 43.9%; Score 38.2; DB 10; Length 439;
Matches 163; Conservative 0; Mismatches 208; Indels 0; Gaps 0;

QY 340 ATAGCAAGACTGCTACTGAGTTGATTAACCGAATGACGCTTATTAAAGTTCTTGATA 399
DB 374 ATGGTGATGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 315
QY 400 CCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 459
DB 314 ATGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 255
QY 460 TTAGAAAAGTGTTCCTCCGGGGAAGTTCGACGCTCCAGATGATGATGATGATGAT 519
DB 254 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 195
QY 520 ATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 579
DB 194 GTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 135
QY 580 AAGATTCAACACTTGTAAAGTAGTATGATGATGATGATGATGATGATGATGATGAT 639
DB 134 GTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 75
QY 640 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 699
DB 74 ATAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 15
QY 700 TTGGTGAGAT 710
DB 14 ATGGTGATGAT 4

RESULT 4
US-09-816-095-3/C
Sequence 3, Application US/09816095
Patent No. US20020137164A1
GENERAL INFORMATION:
APPLICANT: GAN, Weiniu
TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
TITLE OF INVENTION: THEROPF
FILE REFERENCE: C1001147
CURRENT APPLICATION NUMBER: US/09/816.095
CURRENT FILING DATE: 2001-03-26
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 99916
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
```

```
LOCATION: (1)...(99916)
OTHER INFORMATION: n = A,T,C or G
US-09-816-095-3

Query Match
Best Local Similarity 46.2%; Score 36; DB 10; Length 99916;
Matches 120; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 425 TTTTGTGGCTTGGATATTGAGTGAGACCAAGTTTGAAGAAAGTGTCTCCCGGGAA 484
DB 41649 TTCTATTGATGTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 41590
QY 485 GGTGGGACTGCTCAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 544
DB 41589 GAATTTGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 41530
QY 545 TTCTGATATCCCAAGTCTCCACATGCTTATGAAATTCACACTGTAAAGTAG 604
DB 41529 TGCCCTAATTTTCAGAAAGCTGCTCAGTCACTAATTTAGAGATTCACCTTCTCCGTTTAA 41470
QY 605 TATTGGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 664
DB 41469 TCTTTGGAGGGGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 41410
QY 665 TGTGAGATCTTTCAGATT 684
DB 41409 TCGGTAGAGGTGTTTATGAT 41390

RESULT 5
US-09-771-208-20
Sequence 20, Application US/09771208
Patent No. US2002015564A1
GENERAL INFORMATION:
APPLICANT: MEDPARA, JUAN
APPLICANT: BRADFORD, ERIC
APPLICANT: HORVAT, SIMON
TITLE OF INVENTION: CLONING OF A HIGH-GROWTH GENE
FILE REFERENCE: 407P-923710US
CURRENT APPLICATION NUMBER: US/09/771.208
CURRENT FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: US 08/999,477
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.0
SEQ ID NO 20
LENGTH: 659158
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: misc_feature
LOCATION: (123459)...(123478)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (602466)...(602485)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (546998)...(547017)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (494715)...(494814)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (390986)...(391005)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (346860)...(346823)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (317174)...(317193)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (280353)...(280373)
```



```

: NAME/KEY: misc_feature
: LOCATION: 843, 852, 855, 861, 864, 876, 879, 882, 888, 894, 897, 900
: LOCATION: 903, 906, 909, 921, 924, 939, 948
: OTHER INFORMATION: n = A,T,C or G
: US-09-947-971-3

```

Query Match	3.38; Score 35; DB 10; Length 954;
-------------	------------------------------------

[illegible]

```

1 RESULT 9
2 US-09-933-797-413/C
3 : Sequence 413, Application US/09933797
4 : Patent No. US2002015119A1
5 : GENERAL INFORMATION:
6 : APPLICANT: Robert A. Sikes et al.
7 : TITLE OF INVENTION: Isolation and Use of
8 : TITLE OF INVENTION: Sinus Expressed Sequ
9 : FILE REFERENCE: 9901-007-899
10 : CURRENT APPLICATION NUMBER: US/09/933,797
11 : CURRENT FILING DATE: 2001-08-22
12 : PRIOR APPLICATION NUMBER: US/09/482,933
13 : PRIOR FILING DATE: 2000-01-14
14 : PRIOR APPLICATION NUMBER: PCF/US99/10746
15 : PRIOR FILING DATE: 1999-05/14
16 : PRIOR APPLICATION NUMBER: 60/085,383
17 : PRIOR FILING DATE: 1998-05-14
18 : NUMBER OF SEQ ID NOS: 811
19 : SOFTWARE: FastSeq for Windows Version 3.0
20 : SEQ ID NO 413
21 : LENGTH: 280
22 : TYPE: DNA
23 : ORGANISM: Murine
24 : US-09-933-797-413

```

	Query Match	3.3%	Score 34.6	DB 9	Length 280
	Best Local Similarity	49.7%	Pred. No. 1,2		
	Matches	85	Conservative	1	Mismatches 85; Indels 0; Gaps 0
QY	14	TTTTTATTTT	TTTTGTTTACGTAAAGAATGTCATCGTCAAAATTTGATCGACAGCCGTTT	73	
Db	182	TCTCTTTTGTCTTTTGAATTCGCGGCTGGMGATTCCTAACTACCACATTTTGAATGTTT	123		
XY	74	TACAGAGGAAGAGCTTCCTCGCTATATGAGCGCCATCGAAGAGTTCCTACAAATTTTCGCCGTTTC	133		

Db	122	AACGACACAGAAATCTATCTTACTGTCCCTTGAAGTCAATGGCTTCACAGATC	65
Oy	134	TTCTTTCTTCCTTCCTGCTGCTGCACGCTACAGCTCAACCTCCG	184
Db	62	TTTCCCTGTGTTTGGCTCCATTTTCCTTTCCAAATTACCAACATCTCCGG	12

```

RESULT 10
: Sequence 5350, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
:   APPLICANT: Penn, Sharon G.
:   APPLICANT: Rank, David R.
:   APPLICANT: Hanzel, David K.
:   APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: FILE REFERENCE: Aesomica-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: PRIORITY FILING DATE: 2001-05-23
: PRIORITY APPLICATION NUMBER: US 60/180,312
: PRIORITY FILING DATE: 2000-02-04
: PRIORITY APPLICATION NUMBER: US 60/207,456
: PRIORITY FILING DATE: 2000-05-26
: PRIORITY APPLICATION NUMBER: US 09/632,366
: PRIORITY FILING DATE: 2000-08-03
: PRIORITY APPLICATION NUMBER: GB 24263.6
: PRIORITY FILING DATE: 2000-10-04
: PRIORITY APPLICATION NUMBER: US 60/236,359
: PRIORITY FILING DATE: 2000-09-27
: PRIORITY APPLICATION NUMBER: PCT/US01/00666
: PRIORITY FILING DATE: 2001-01-30
: PRIORITY APPLICATION NUMBER: PCT/US01/00667
: PRIORITY FILING DATE: 2001-01-30
: PRIORITY APPLICATION NUMBER: PCT/US01/00664
: PRIORITY FILING DATE: 2001-01-30
: PRIORITY APPLICATION NUMBER: PCT/US01/00669
: PRIORITY FILING DATE: 2001-01-30
: PRIORITY APPLICATION NUMBER: PCT/US01/00665
: PRIORITY FILING DATE: 2001-01-30
: PRIORITY APPLICATION NUMBER: PCT/US01/00668
: PRIORITY FILING DATE: 2001-01-30
: PRIORITY APPLICATION NUMBER: PCT/US01/00663
: PRIORITY FILING DATE: 2001-01-30
: PRIORITY APPLICATION NUMBER: PCT/US01/00662
: PRIORITY FILING DATE: 2001-01-30
: PRIORITY APPLICATION NUMBER: PCT/US01/00661
: PRIORITY FILING DATE: 2001-01-30
: PRIORITY APPLICATION NUMBER: PCT/US01/00670
: PRIORITY FILING DATE: 2001-01-30
: PRIORITY APPLICATION NUMBER: US 60/234,687
: PRIORITY FILING DATE: 2000-09-21
: PRIORITY APPLICATION NUMBER: US 09/608,408
: PRIORITY FILING DATE: 2000-06-30
: PRIORITY APPLICATION NUMBER: US 09/774,203
: PRIORITY FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 5350
:   LENGTH: 463
:   TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC000053.1
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.8
: OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2

```

RESULT 13
US-09-864-761-8066/c
; Sequence 8066, Application US/09864761
; Patent No. US20020048763A1

```

; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 8066
; LENGTH: 527
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO ACO13707.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.2
; US-09-864-761-8066

Query Match          3.2%; Score 34; DB 10; Length 527;
Best Local Similarity 50.6%; Pred. No. 2.7;
Matches 82; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

Qy 531 ATGCAATATTTTCATCTGCTATCCCTCAAGATCTCCACACTTTATGGAATTCACA 590
    || || || || || || || || || || || || || || || || || || || ||
Db 226 ATCCACATTTTATGTGTGGAAGATATATAGCTATATATATATATATTCACCAAGTAA 167
Qy 591 CTTTGAAGGTAGCTATTGGAATTTGATGTGACTCTGTGAAGCTTTTCCATGACTATGGA 650
```

```

; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30

RESULT 14
US-09-925-302-270/c
; Sequence 270; Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 270
; LENGTH: 1827
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-302-270

Query Match          3.2%; Score 34; DB 10; Length 1827;
Best Local Similarity 56.1%; Pred. No. 5.7;
Matches 64; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Qy 10 TTAATTTTATTTTTTGTTCAGTAAAGAAATGTCATGCAATTCGAATTCGATGACGAG 69
    || || || || || || || || || || || || || || || || || || || ||
Db 1808 TTTTATTTTATTTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1749
Qy 70 CTTTACAGAGGAAGACTTCTGCTATGCGCCGCAATCGAAGCTTCTCAATTT 123
    || || || || || || || || || || || || || || || || || || || ||
Db 1748 ACTGTGCCCGCAAGTGTCTCCCAACATGSAACCAAGAACCAAGACAGTAAATTT 1695

RESULT 15
US-09-864-761-8619
; Sequence 8619; Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
```

```

? PRIOR APPLICATION NUMBER: PCT/US01/00669
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00665
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00668
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00663
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00662
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00661
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00670
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: US 60/234,687
? PRIOR FILING DATE: 2000-09-21
? PRIOR APPLICATION NUMBER: US 09/608,408
? PRIOR FILING DATE: 2000-06-30
? PRIOR APPLICATION NUMBER: US 09/774,203
? PRIOR FILING DATE: 2001-01-29
? NUMBER OF SEQ ID NOS: 49117
? SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
? SEQ ID NO 8619
? LENGTH: 556
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? OTHER INFORMATION: MAP TO AC023000.2
? OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
? OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
? OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8
? OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
? OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
US-09-864-761-8619

```

Query Match	3.2%	Score	33.8	DB	10	Length	556
Best Local Similarity	46.7%	Pred. No.	3.2				
Matches 107; Conservative	0	Mismatches	122			Indels	0
						Gaps	0

Qy	498	CAGCATATCTGAGAACTAAATATATTTGGATGCTTTCGATATTTTCTCTGGTATCCCT	557
	111	111	111
Db	164	CATATATGCTGATGATGATGGTGGTCTTATGCGCATATATGTGATGATGGTGA	223
	111	111	111
Qy	558	CAAACTCTCAACATCTTATTTAAGATTCAACACTGTGAAAGTATGGATTTGGAATTTGAT	617
	111	111	111
Db	224	GAT	283
	111	111	111
Qy	618	GCTGACTCTGTGAAGCTTTTCCATGACTATGAGACTTAGTATCAACAGATGTTGAGGATCTT	677
	111	111	111
Db	284	AGCTGATGCTGATGATGCTGATGAATGAATGATGATGCTGATCATGATGATGATGCTGATAGT	343
	111	111	111
Qy	678	TTCAGATTTAGCCCAACCAAAAAATTTGCTGAGATTTAAAAAATGGGGCCCTTG	726
	111	111	111
Db	344	GATGATGATGATGATGCTGATGAATGATGATGATGATGATGATGCTGGGGCCGATG	392
	111	111	111

Search completed: November 2, 2002, 16:42:44
Job time : 788 secs

Db 433 GATATTAGCATGAGTCTATCAGATGGGAGTGGTGGGATTTGACATGAGTGGCCACA 492
Qy 456 AGTTTGAAGAAAGTGTCTCCCGGAGAGTTGCGACTGTCGATATGTGATAGT 515
Db 493 TTATNCATATGAGGGAAGT---TGCGAAGTTGCTACTAATTCAGTGTGTTCTGAG 549
Qy 516 AATTATTGTGATGTTATGATATTTT---CATTTGTGATCCCTCAAGTCTCCAAAT 572
Db 550 AGCAAAATGTTACTGTTCCAGCTTCTTCCATGTCAGTTTTCGCCAGGATTAAGATG 609
Qy 573 CTATTGAAGATTCACACTGTTGAAGTGTGTAAGTGTGTAATGTATGATGCTGTGAAG 632
Db 610 TTGCTTAAATTAAGGATTAAGAAAGGAGGTGTAGAAATGGAAGATCACTGGA 669
Qy 633 CTTTCCATGACTGTGAGTGTAGTATCAAGATGTGAGATCTTTGATAGTTAGCCAC 692
Db 670 CTTTACTGACTGTTGATATCAAAATTTTGTGAGTTGACAGATGTTGCCAAT 729
Qy 693 CAAAAATTTGGTGAGATTAAGAAATGGGCTTGGCTCCTACTACTGAGACACTGTTTGC 752
Db 730 AAAAAGCTGAATGTACAGAGACCTGGACCTTAAACAGTCTGGTTAAACACCTTTAGT 789
Qy 753 AAAGAGTCTCTGAAGCCAAACAGATCAGGCTTGGAACTGGAGTTTATCCTCTGCA 812
Db 790 AAACAGCTCTGAAAGACAGTCTATCCGCTGTAAGCAATTTGAGTAATTTCTCTCACT 849
Qy 813 AAGCAGAGTTTACATACGAGCAACGAGATGCTTATGCTTCACTGATCTTTACAGGT 872
Db 850 GAGACCCAGAAACCTGTATGACAGCCACTGATGCTTATGCTGTTTATTTTACCGAAT 909
Qy 873 CTTAAGGACCTTCCGTATGCTGT 895
Db 910 TTAGAGATTTTGGATGATGACTGT 932

RESULT 2
US-09-791-211-11
; Sequence 11, Application US/09791211
; Patent No. 6448080
; GENERAL INFORMATION:
; APPLICANT: Donna T. Ward
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION
; FILE REFERENCE: RTS-0205
; CURRENT APPLICATION NUMBER: US/09/791, 211
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 90
; SEQ ID NO 11
; LENGTH: 5208
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (232)...(4530)
US-09-791-211-11

Query Match 10.6%; Score 111; DB 4; Length 5208;
Best Local Similarity 54.1%; Pred. No. 2.5e-25;
Matches 272; Conservative 0; Mismatches 225; Indels 6; Gaps 2;

Qy 396 GATACCAAGAGATGATCTGCAATAGCTTTTGGCTTGGATATGATGAGACCA 455
Db 433 GATATTGATGATGATCTGATGATGATGATGATGATGATGATGATGATGATGAT 492
Qy 456 AGTTTGAAGAAAGTGTCTCCCGGAGAGTTGCGACTGTCGATATGTGATAGT 515
Db 493 TTATNCATATGAGGGAAGT---TGCGAAGTTGCTACTAATTCAGTGTGTTCTGAG 549
Qy 516 AATTATTGTGATGTTATGATATTTT---CATTTGTGATCCCTCAAGTCTCCAAAT 572
Db 550 AGCAAAATGTTACTGTTCCAGCTTCTTCCATGTCAGTTTTCGCCAGGATTAAGATG 609
Qy 573 CTATTGAAGATTCACACTGTTGAAGTGTGTAAGTGTGTAATGTATGATGCTGTGAAG 632

Db 610 TTGCTTGAATTAAGCAGTTAAAGGAGGAGTGTAGGAATTAAGGAGATTCAGTGGAAA 669
Qy 633 CTTTCCATGACTATGAGTGTAGTATCAAGATGTGAGATCTTTCAGTATTAAGCAAC 692
Db 670 CTTTACTGACTGTTGATATCAAAATTTTGTGAGTTGACAGATGTTGCCAAT 729
Qy 693 CAAAAATTTGGTGAGATTAAGAAATGGGCTTGGCTCCTACTACTGAGACACTGTTTGC 752
Db 730 AAAAAGCTGAATGTACAGAGACCTGGACCTTAAACAGTCTGGTTAAACACCTTTAGT 789
Qy 753 AAAGAGTCTCTGAAGCCAAACAGATCAGGCTTGGAACTGGAGATTTATCCTCTGCA 812
Db 790 AAACAGCTCTGAAAGACAGTCTATCCGCTGTAAGCAATTTGAGTAATTTCTCTCACT 849
Qy 813 AAGCAGAGTTTACATACGAGCAACGAGATGCTTATGCTTCACTGATCTTTACAGGT 872
Db 850 GAGACCCAGAAACCTGTATGACAGCCACTGATGCTTATGCTGTTTATTTTACCGAAT 909
Qy 873 CTTAAGGACCTTCCGTATGCTGT 895
Db 910 TTAGAGATTTTGGATGATGACTGT 932

RESULT 3
US-09-127-670-5
; Sequence 5, Application US/09127670
; Patent No. 6228583
; GENERAL INFORMATION:
; APPLICANT: Massachusetts Institute for Technology
; APPLICANT: Leonard P. Guarente
; APPLICANT: David A. Sinclair
; APPLICANT: David B. Lombard
; TITLE OF INVENTION: ASSAYS FOR COMPOUNDS WHICH EXTEND LIFE
; FILE REFERENCE: MIT-7720DA
; CURRENT APPLICATION NUMBER: US/09/127, 670
; CURRENT FILING DATE: 1998-07-31
; EARLIER APPLICATION NUMBER: US 60/054, 629
; EARLIER FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 6476
; TYPE: DNA
; ORGANISM: Murine
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (229)...(4432)
US-09-127-670-5

Query Match 7.9%; Score 82.8; DB 4; Length 6476;
Best Local Similarity 53.2%; Pred. No. 3.7e-16;
Matches 199; Conservative 0; Mismatches 172; Indels 3; Gaps 1;

Qy 481 GGAAGTTGGAGCTGTCAGATATGTGTAGATGATTAATTTGATGATGATGATGAT 540
Db 494 GCAAGAGCGAGATGATCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 553
Qy 541 TT---CATTTGTGATCCCTCAAGTCTCAACATCTTATGGAAGATTCACACTGTGA 597
Db 554 CTTCCATGATGATGTTTCCCGGAGGATTAAGAAATGTTACTAGAAACAAATTCATTAAGA 613
Qy 598 AGGATGATTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 657
Db 614 AGGACGAGGATTTGAGATTTGAAGGAGGACGAGAACTTCTGCTGATTTTGAAGTCAAGT 673
Qy 658 TCAAGATGTTGAGATCTTTCAGATTTTACCAACCAAAAAATTTGATGATGATGATGAT 717
Db 674 TGAGAGATTTTGTGAGAGCTACGAGATGTTTCCAAATGGAAGTTGAAGTGCAGAGACT 733
Qy 718 GGGGCTTGGCTCCTACCTGAGACACTGTTTGGCAAGAGCTCTGGAAGCCAAACAGAA 777

Db 734 GGAGCCATATGCTGCTTAACAGCTTTAGGGAACACTTTTGAAAGACAGTCCA 793
QY 778 TCAGGCTTGGGAAGTGGAGCTTTATCTCTGCAAGACAGCAGTTCAAATGACAGCAA 837
Db 794 TCCGCTGACAGCAATGGAGTAATTTCCCTCACTAGAGACAGAAATCTATGACAGCCA 853
QY 838 CGAGTCTTATGCT 851
Db 854 CTGATGCTTATGCT 867

RESULT 4

US-08-781-891-205
; Sequence 205, Application US/08781891
; Patent No. 6090620
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Yu, Chang-En
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; APPLICANT: Schellenberg, Gerald D.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; TITLE OF INVENTION: WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,891
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6090620tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 240052.419
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 205:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4792 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 145..4347
; US-08-781-891-205

Query Match 7.7%; Score 81.2; DB 3; Length 4792;

Best Local Similarity 52.3%; Pred. No. 1e-15; Indels 0; Gaps 0;
Matches 179; Conservative 0; Mismatches 163;

QY 510 GATAGTAATATGTGATGTTATGATATTTTCATTTGATGCTTCCCTCAAGCTTCACAA 569
Db 442 GAGAGCAATGTTACTGTTTCACATTTCTTCATGTCAGTTTCCCGGAGATTAA 501
QY 570 CATCTATTGAGATTCAACACTTGTAAAGGTAGTATGGAATGATGCTGCTG 629
Db 502 ATGTTACTAGAAACAAATCAATTAAGAGGCGAGGTTGGATGGAAGCGACAGTGG 561
QY 630 AACCTTTTCATGATGAGATTAGTATGCAAGATGCTTGAGATCTTTCAGATTAGCC 689
Db 562 AAACCTTCGCTGATTTTGACGTCAAGTTGGAGAGTTTGTGAGAGTGCAGATGTTGCC 621

QY 690 AACCAAAAATGGTGGAGATAAATAATGGGCTTGCCTCACTAATGAGACATTTGTT 749
Db 622 AATGAAAAGTTGAAGTGCAGAGACCTGAGCCTCAATGATCTGTGTTAAACAGCTTAA 681
QY 750 TCCAAAGAGCTCCTGAAGCCAAACAGAAATCAGCTTGGAGACTGGAGTTTATCTCTG 809
Db 682 GGGAAACACCTTTTGAAAGCAAGTCATCCGCTGACAGCAATTGGAGTAATTTCCCTC 741
QY 810 TCAAGCAGAGTTACATACGAGCAACGATGCTTATGCT 851
Db 742 ACTGAGGACCAAGAACTGTATGCAAGCCTGATGCTTATGCT 783

RESULT 5

US-09-791-211-10
; Sequence 10, Application US/09791211
; Patent No. 6448080
; GENERAL INFORMATION:
; APPLICANT: Donna T. Ward
; APPLICANT: Andrew T. Walt
; TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION
; FILE REFERENCE: RTS-0205
; CURRENT APPLICATION NUMBER: US/09/791,211
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 90
; SEQ ID NO 10
; LENGTH: 98844
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 24962
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 64383
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 65468
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 65469
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 65470
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 65471
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 87130
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 89049
; OTHER INFORMATION: unknown
; OTHER INFORMATION:
; US-09-791-211-10

Query Match 5.3%; Score 55.4; DB 4; Length 98844;

Best Local Similarity 60.1%; Pred. No. 1.4e-06; Indels 0; Gaps 0;
Matches 92; Conservative 0; Mismatches 61;

QY 545 TTCTGATATCCCTCAAGATCCCAACATCTTATTGAAGATTCAACCTGTAAGGTAGG 604
Db 45421 TACAGTTTTCCCGAGGATTAATAATGTGCTGTAATAAATGAAGCACTTAAAGGAGG 45480
QY 605 TATTGGAATTGATGAGTCTGTGAAGCTTTTCATGACTATGAGATTAGTCAAGA 664
Db 45481 TGTAGGAATGGAAGGATGATGGAAGCAATCTTACGTGACTTGTATATCAAAATTGAAGA 45540
QY 665 TGTGAGAGATCTTTCAGATTATGACCAACCAAA 697
Db 45541 TTTTGGAGCTTGACAGATGTTGCCAATAAAA 45573

RESULT 6

US-08-781-891-207

Sequence 207, Application US/08781891
Patent No. 6090620

GENERAL INFORMATION:

APPLICANT: Fu, Ying-Hui
APPLICANT: Yu, Chang-En
APPLICANT: Oshima, Junko
APPLICANT: Mulligan, John T.
APPLICANT: Schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
TITLE OF INVENTION: WERNER'S SYNDROME
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,891
FILING DATE: 27-DEC-1996

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:
NAME: No. 6090620Leuburg Ph.D., Carol

REGISTRATION NUMBER: 39,317

REFERENCE/DOCKET NUMBER: 240052.419

TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 207:

SEQUENCE CHARACTERISTICS:

LENGTH: 29604 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-781-891-207

Query Match 4.9%; Score 51.2; DB 3; Length 29604;
Best Local Similarity 56.5%; Pred. No. 1.5e-05;

Matches 95; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 545 TCTGTATCCCTCAAGTCTCCACATCTTATGAGATTCAACCTGTAAAGGTAGG 604

DB 19789 TTTAGTTTCCCTCCAGGAGTAAATGTTACTAGAAACAAATCAATTAAGAGCAGG 19848

QY 605 TATTTGAAATGAGTGACCTCTGTGAAGCTTTCCATGACTATGAGCTTAGTATCAAGA 664

DB 19849 GGTGTGAGATTGAAGGGGAGCAGTGAACCTTCTGCGTATTGACCTCAAGTTGGAGAG 19908

QY 665 TCTTTGAGATCTTTCAGATTACCAACCAAAAATTTGGTGAGATTA 712

DB 19909 TTTTGTGAGACTGACGAGATCTTCCCAATGAAAAGTAGGCGTAATATA 19956

RESULT 7

US-08-232-463-14

Sequence 14, Application US/08232463
Patent No. 5670367

GENERAL INFORMATION:

APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/935,313

FILING DATE:

APPLICATION NUMBER: EP 91 114 300.6

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 30472/114 IMMU

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300

TELEFAX: (703)683-4109

TELEX: 899149

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 7218 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

CLONE: pTZgpt-fls

US-08-232-463-14

Query Match 4.7%; Score 49.2; DB 1; Length 7218;
Best Local Similarity 11.5%; Pred. No. 2.8e-05;

Matches 33; Conservative 140; Mismatches 113; Indels 0; Gaps 0;

QY 9 ATTAATTTTATTTTCTTCTTCAGTAAAGAAATGTCATGCTCAATTTGATCGACGAC 68

DB 992 ACTATTTTCTTCCTTCCTTCATGCTTCACAGATTAATTAATTCGAGCTGCGAGGTC 1051

QY 69 GCTTTACAGAGAGAGAGCTTCGCTATCGACGCGCATGGAAGCTTCCATATTTCTCC 128

DB 1052 GAGGAGCTTGCATATTT 1111

QY 129 CGTTCTTCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 188

DB 1112 YY 1171

QY 189 GCGCAGAGAGAGAGATCAATCAATCCCAATATATATTCGTCGCAATGCTGCTGTC 248

DB 1172 YY 1231

QY 249 ATCACTTCTTCATCTATTAAGCATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 294

DB 1232 YY 1277

RESULT 8

US-09-300-672-3

Sequence 3, Application US/09300672
Patent No. 6248937

GENERAL INFORMATION:

APPLICANT: FINKELSTEIN, Ruth R.
APPLICANT: Lynch, Tim
APPLICANT: Goodman, Howard M.
APPLICANT: Wang, Ming-Li

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 31, 2002, 13:38:19 ; Search time 27 Seconds
(without alignments)
1845.279 Million cell updates/sec

Title: US-09-896-186b-24
Perfect score: 1491
Sequence: 1 MSSSNWIDDAFTEELDAID.....YASWMLKYKLDLPDAVSGS 288

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1466.5	98.4	285	10	Q9FT68
2	1439.5	96.5	313	10	Q9SVM6
3	339.5	22.8	1436	13	Q9S330
4	339	22.7	968	11	Q9Z241
5	339	22.7	1401	11	Q9JKD4
6	330.5	22.2	643	11	Q9S948
7	238	16.0	201	10	Q93VU9
8	229.5	15.4	346	5	Q9VE86
9	221.5	14.9	353	5	Q961E1
10	212.5	14.3	583	5	Q9VGN7
11	179.5	12.0	582	10	Q9C7K6
12	173.5	11.6	494	10	Q93VS2
13	159.5	10.7	699	3	Q9C2I6
14	158	10.6	123	10	Q22964
15	151.5	10.2	199	10	Q9LK79
16	151	10.1	496	4	Q9NVH0

17	142	9.5	599	4	Q96NP1	Q96NP1 homo sapien
18	127.5	8.6	239	10	Q9S1H3	Q9S1H3 arabidopsis
19	121	8.1	505	10	Q9FTN8	Q9FTN8 arabidopsis
20	112.5	7.5	217	10	Q80886	Q80886 arabidopsis
21	111.5	7.5	210	10	Q9LHG3	Q9LHG3 arabidopsis
22	111	7.4	242	10	Q9LHG2	Q9LHG2 arabidopsis
23	111	7.4	300	2	Q9AL74	Q9AL74 citrobacter
24	110	7.4	238	10	Q9C7A5	Q9C7A5 arabidopsis
25	105	7.0	625	5	Q9VTE1	Q9VTE1 drosophila
26	103	6.9	220	10	Q9LHG1	Q9LHG1 arabidopsis
27	102.5	6.9	710	5	Q9NA69	Q9NA69 caenorhabdi
28	101	6.8	876	5	Q17951	Q17951 caenorhabdi
29	98.5	6.6	265	10	Q9LHG5	Q9LHG5 arabidopsis
30	98.5	6.6	532	17	Q976Y3	Q976Y3 sulfolobus
31	97.5	6.5	445	17	Q9YDV0	Q9YDV0 aeropyrum P
32	97	6.5	579	2	Q9ROK2	Q9ROK2 bacillus st
33	97	6.5	906	5	Q17281	Q17281 botryllus s
34	96.5	6.5	488	16	Q97JZ0	Q97JZ0 clostridium
35	95	6.4	455	10	Q9ZWS2	Q9ZWS2 vigna mungo
36	95	6.4	925	5	Q9GUG1	Q9GUG1 caenorhabdi
37	95	6.4	4589	5	Q76506	Q76506 tetrahymena
38	94.5	6.3	353	10	Q9LHG4	Q9LHG4 arabidopsis
39	94.5	6.3	416	5	Q96144	Q96144 plasmodium
40	93	6.2	488	17	Q9HJ76	Q9HJ76 thermoplasma
41	93	6.2	552	10	Q9LVM8	Q9LVM8 arabidopsis
42	93	6.2	2224	4	Q43737	Q43737 homo sapien
43	92.5	6.2	721	10	Q94B18	Q94B18 capsicum an
44	92.5	6.2	1097	3	Q13592	Q13592 saccharomyc
45	92	6.2	237	10	Q9ZV70	Q9ZV70 arabidopsis

ALIGNMENTS

RESULT 1

Query Match 98.4%; Score 1466.5; DB 10; Length 285;
Best local similarity 99.0%; Pred. No. 1.2e-126;
Matches 285; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

Query 1 MSSSNWIDDAFTEELDAIDAEASYNFSRSSSSSSAAPTVAQTTSVHGHEEDNPIN 60
DB 1 MSSSNWIDDAFTEELDAIDAEASYNFSRSSSSSSAAPTVAQTTSVHGHEEDNPIN 60
QY 61 NIREQLPRSTSTSYKRFPLSRANRPFAMRGGRILYKTEVDKRAMOLIKYLD 120
DB 61 NIREQLPRSTSTSYKRFPLSRANRPFAMRGGRILYKTEVDKRAMOLIKYLD 120
QY 121 KRDEGIAFVGLDIEMRPSPKGVLPKGVATVQICVDSNYCDVMHIFRSGIPSLQHLIE 180

```
Db 121 KRDESGIAFGLDIEMRSPFRKGVLPKGVATVQICVDSNYCDVMHIFHSGIPQSLQHLIE 180
Qy 181 DSTLVKVGIGIDGSVKLFHDYGVSIKDVEDLSLANOKIGGDKKMWGLASTETLVCKEL 240
Db 181 DSTLVKVGIGIDGSVKLFHDYGVSIKDVEDLSLANOKIGGDKKMWGLASTETLVCKEL 240
Qy 241 LKPNRIRLGNWFFYPLSKQOLQVATADAYASWMLYKVLKDLDPDAVSGS 288
Db 241 LKPNRIRLGNWFFYPLSKQOLQVATADAYASWMLYKVLKDLDPDAVSGS 285

RESULT 2
Q9SYV6 PRELIMINARY; PRT; 313 AA.
ID Q9SYV6
AC Q9SYV6;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE HYPOTHEICAL 35.1 KDA PROTEIN.
GN F18A5.260 OR AT4G13870.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
ON NCBI_TaxID=3702.
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Weber N., Grueninger D., Schmidheini T., Bancroft I.,
RA Mewes H.W., Mayer K.F.X., Lemcke K., Schueller C.,
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Weber N., Grueninger D., Schmidheini T., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL035528; CAB36851.1; -
DR EMBL; AL161537; CAB78429.1; -
DR InterPro: IPR002562; 3.5.exonuclease.
DR Pfam: PF01612; 3.5.exonuclease; 1.
DR SMART; SM00474; 3EXOC; 1.
DR Hypothetical protein.
SQ SEQUENCE 313 AA; 35105 MW; 448D7ACC375D4D22 CRC64;

Query Match 96.5%; Score 1439.5; DB 10; Length 313;
Best Local Similarity 95.3%; Pred. No. 4e-124; 4; Indels 9; Gaps 1;
Matches 281; Conservative 1; Mismatches 281;

Qy 1 MSSSNWIDDAFTBEELLALDAIEASYNFSRSSSSSAAPLVQATTSVHGHEEDPNQIPN 60
Db 1 MSSSNWIDDAFTBEELLALDAIEASYNFSRSSSSSAAPLVQATTSVHGHEEDPNQIPN 60
Qy 61 NNRQLPRSTSTSYKRFPLSRANFRANFRGGRILYKATETVDKRAMOLIKVLDLT 120
Db 61 NNRQLPRSTSTSYKRFPLSRANFRANFRGGRILYKATETVDKRAMOLIKVLDLT 120
Qy 121 KRDESGIAFGLDIEMRSPFRKGVLPKGVATVQICVDSNYCDVMHIFHSGIPQSLQHLIE 180
Db 121 KRDESGIAFGLDIEMRSPFRKGVLPKGVATVQICVDSNYCDVMHIFHSGIPQSLQHLIE 180
Qy 181 DSTLVKVGIGIDGSVKLFHDYGVSIKDVEDLSLANOKIGGDKKMWGLASTETLVCKEL 240
Db 181 DSTLVKVGIGIDGSVKLFHDYGVSIKDVEDLSLANOKIGGDKKMWGLASTETLVCKEL 240
Qy 241 LKPNRIRLGNWFFYPLSKQOLQVATADAYASWMLYKVLKDLDPDAVSGS 286
```

```
Db 241 LKPNRIRLGNWFFYPLSKQOLQVATADAYASWMLYKVTYTKNHLITLNDLEAKIS 295

RESULT 3
Q93530 PRELIMINARY; PRT; 1436 AA.
ID Q93530
AC Q93530;
DT 01-NOV-1998 (TREMblrel. 08, Created)
DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)
DE WERNER SYNDROME HELICASE HOMOLOG (FOCUS FORMING ACTIVITY 1).
GN FFA-1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Piploidea; Pipidae;
OC Xenopodinae; Xenopus.
ON NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A. AND FUNCTION.
RX MEDLINE=98361165; Pubmed=9697700;
RA Van H., Chen C.-Y., Kobayashi R., Newport J.;
RT "replication focus-forming activity 1 and the werner syndrome gene
RL Nat. Genet. 19:375-378(1998).
RN [2]
RP FUNCTION.
RX MEDLINE=96032793; Pubmed=7569932;
RA Van H., Newport J.;
RT "FFA-1, a protein that promotes the formation of replication centers
RT within nuclei."
RL Science 269:1883-1885(1995).
CC -1- FUNCTION: ESSENTIAL FOR THE FORMATION OF DNA REPLICATION FOCAL
CC CENTERS: STABLY ASSOCIATES WITH FOCI ELEMENTS GENERATING BINDING
CC SITES FOR RP-A. EXHIBITS A MAGNESIUM DEPENDENT ATP-DEPENDENT DNA-
CC HELICASE ACTIVITY. MAY BE INVOLVED IN THE CONTROL OF GENOMIC
CC STABILITY.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE RECQ SUBFAMILY OF HELICASES.
CC -1- SIMILARITY: CONTAINS 1 HRDC DOMAIN.
DR EMBL; AF067418; AAC63512.1; -
DR InterPro: IPR002562; 3_5_exonuclease.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR001650; Helicase_C.
DR InterPro: IPR002121; HRDC.
DR Pfam: PF01612; 3.5.exonuclease; 1.
DR Pfam: PF00271; DEAD; 1.
DR Pfam: PF00271; Helicase_C; 1.
DR Pfam: PF00570; HRDC; 1.
DR SMART; SM00487; 3EXOC; 1.
DR SMART; SM00490; HELIC; 1.
DR SMART; SM00341; HRDC; 1.
DR Hydroxylase; Helicase; ATP-binding; DNA-binding; Nuclear protein.
KW DOMAIN 461 467
FT DOMAIN 1098 1117 HRDC.
FT NE_BIND 512 519 ATP (BY SIMILARITY).
FT SITE 609 612 DEAD BOX.
SQ SEQUENCE 1436 AA; 161850 MW; 1BEAF05A25BAE230 CRC64;

Query Match 22.8%; Score 339.5; DB 13; Length 1436;
Best Local Similarity 35.0%; Pred. No. 4.2e-22;
Matches 82; Conservative 43; Mismatches 84; Indels 25; Gaps 7;

Qy 61 NNRQLPRSTSTSYKRFPLSRANFRANFRGGRILYKATETVDKRAMOLIKVLDLT 108
Db 3 SLOKRLPEMWSVYQOEDRIDDAKKSFCCKNILEDNLPFMKFNCSIYVSYESNCCSLSED 62
Qy 109 KRAMOLIKVLDTRKDESGIAFGLDIEMRSPFRKGVLPKGVATVQICVDSNYCDVMHIF- 167
Db 63 IRRSL------EEDV--LGFIDIEPPIYTKG-KTGVALIQLQVSSKKCYLFIISP 110
Qy 168 HSGIPQSLQHLIEDSTLVKVGIGIDGSVKLFHDYGVSIKDVEDLSLANOKIGGDKKMW 227
```



```

Oy 280 DL 281
Db 227 DL 228

RESULT 11
ID 09C7K6 PRELIMINARY; PRT; 582 AA.
AC 09C7K6:
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TEMBLrel. 18, Last annotation update)
DE 3'-5' EXONUCLEASE, PUTATIVE.
GN F14G9.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucrids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Cressy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maill R., Marziani A.,
RA Miltischer J., Miranda M., Nguyen M., Niemman W.C., Osborne B.I.,
RA Sakano H., Salazar S.L., Schwartz J.R., Shinn P., Soultwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vayberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana."
RL Nature 408:816-820(2000).
DR EMBL: AC069159; AAG50917.1;
DR InterPro: IPR002562; 3_5-exonuclease.
DR Pfam: PF01612; 3_5-exonuclease; 1.
DR SMART: SM00474; 35EXOC; 1.
KW Exonuclease.
SQ SEQUENCE 582 AA; 66467 MW; 34D752473E786680 CRC64;

Query Match 12.0%; Score 179.5; DB 10; Length 582;
Best Local Similarity 28.2%; Pred. No. 6.2e-08;
Matches 53; Conservative 36; Mismatches 66; Indels 33; Gaps 7;

Oy 126 GIAFVGLDIEMRSFKGVLPKVAIVQICVDSN--YCDVMIHFHSG--IPQSLQHLIE 180
Db 377 GCRVGVIDCEWKPNYIKSKONKVSIMQIGSDTKIFILDIKILYNASEILDNCLSHILQ 436
Oy 181 DSTLVVY-----GIGIDGSVKLFHDYGSIKDVE-----DLSLDLNOKI 220
Db 437 SKETLVLVSTEDYPPHKSSGYNFOCDIKQLALSTG-DKCEERYDMLIDIONVNEEP 495
Oy 221 GGDKKWGLASLTETLVCKELLPNRIKGNMEFYPLSKQOLQYANTDAYASMLLYKVLKD 280
Db 496 G-----GLAGLTIKILIGVSLNKR--RNSDMQRPLSQNLERALDAANLVLIHFNRVD 548
Oy 281 LPDAVSGS 288
Db 549 HPHDSSS 556

RESULT 12
O93VS2
O93VS2A

```

```

ID 093VS2 PRELIMINARY; PRT; 494 AA.
AC 093VS2:
DT 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE P0487H02.21 OR P0682B08.2 PROTEIN (P0682B08.2 PROTEIN).
GN P0487H02.21 OR P0682B08.2.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponebare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0487H02."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponebare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0682B08."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP002883; BAB67861.1;
DR EMBL: AP003578; BAB60932.1;
SQ SEQUENCE 494 AA; 51497 MW; 8F83F79405C2BD1E CRC64;

Query Match 11.6%; Score 173.5; DB 10; Length 494;
Best Local Similarity 24.3%; Pred. No. 1.8e-07;
Matches 69; Conservative 42; Mismatches 126; Indels 47; Gaps 11;

Oy 13 EEEELLIDALEAYNRSRSSSSSSAAPTQATTSVHG-----HEEDN 56
Db 204 DEELK-----SAGRRARPPSSMGVGEATEFTGIGGCPAPATSSSSPPPPHPSPH 259
Oy 57 QINNNIRQLRSLRSTSTSYKRRPLSRCRARNPARRFGRLYSTATVEVDRRAQLIK 116
Db 260 HTHHFFDRSDIRSA-----PLLDVDGGCTVSFSGSAI-DTIVTSDAAANAEVWR 309
Oy 117 -----VLDTRKDESGIAFNGLDIEMRP--SEKGVLPKVAIVQICVDSNYCDVMIHFS 169
Db 310 RVRASATTPRGCGGL-LVGLDCEMKPCDHLMPAVAP-TVAIIQLCACGS-CLIIQLLHV 366
Oy 170 G-----IPQSLQHLIEPSTLVKVGIGDGSVKLFHDYGSIKNVEPLSLANOKIG--G 222
Db 367 AGARRVPLVGLDLADPSVRLVIGIGENNAKLADGIVGCAAPVLEVDVCDRLRLP 426
Oy 223 DKRWGLASLTETLVCKELLPNRIKGNMEFYPLSKQOLQYANT 266
Db 427 ARRLGLKGYREVLTGLTMEKPMQVTRSDWERRRLDAAGVRYACS 470

RESULT 13
ID 09C2I6 PRELIMINARY; PRT; 699 AA.
AC 09C2I6:
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE RELATED TO WERNER SYNDROME HELICASE.
GN 93G11.60.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN (1)
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nykretura G., Mewes H.W., Mannhaupt G.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 31, 2002, 13:38:04 ; Search time 13 Seconds

(without alignments)
857.787 Million cell updates/sec

Title: US-09-896-186b-24

Perfect score: 1491
Sequence: 1 MSSSNWIDDAFTEELLAID.....YASWHLVKYKLDPAVSGS 288

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query	Length	ID	Description
1	352	23.6	1432	1 WRN_HUMAN	014191 homo sapien
2	339	22.7	1401	1 WRN_MOUSE	009053 mus musculu
3	136.5	9.2	910	1 Y068_CAEEL	P34607 caenorhabdi
4	95	6.4	508	1 CPVL_RAT	P22443 rattus norv
5	93.5	6.3	1525	1 POLG_STEVM	P09732 s genome po
6	93	6.2	2224	1 FAS_HUMAN	P12259 homo sapien
7	93	6.2	3079	1 IRA2_YEAST	P19158 saccharomyc
8	92	6.2	758	1 LEU2_SCHPO	014289 schizosacch
9	91.5	6.1	473	1 VL2_HPV03	P36744 human papil
10	91	6.1	503	1 DPO1_MOUSE	P28649 mus musculu
11	89.5	6.0	1324	1 CPV1_HAEIN	P43741 haemophilus
12	89.5	6.0	1487	1 MSH6_ARATH	004716 arabidopsis
13	89.5	6.0	1487	1 BLM_DROME	09vg18 drosophila
14	88	5.9	259	1 HK22_BRARE	090461 brachydanio
15	88	5.9	416	1 PROA_VIBCH	09kpe9 vibrio chol
16	88	5.9	784	1 Y063_CAEEL	P34603 caenorhabdi
17	87	5.8	579	1 MTC1_BACST	P43423 bacillus st
18	87	5.8	3137	1 CA36_CHICK	P15989 gallus gall
19	86.5	5.8	346	1 PST5_ECOLI	P06128 escherichia
20	86.5	5.8	470	1 PLSR_CUCSA	039639 cucumis sat
21	86	5.8	552	1 FMR2_APLCA	P08021 aplysia cal
22	85.5	5.7	1240	1 DEOC_MYCPI	P47722 mycoplasma
23	85.5	5.7	1394	1 E75B_DROME	P17672 drosophila
24	85.5	5.7	2298	1 C215_HUMAN	09y3r5 homo sapien
25	84.5	5.7	480	1 UGDH_SOYBN	096558 glycine max
26	84.5	5.7	830	1 GYRA_CLOAB	P94605 clostridium
27	83.5	5.6	777	1 YASB_SCHPO	010146 schizosacch
28	83	5.6	482	1 6PGD_HUMAN	P52209 homo sapien
29	83	5.6	1129	1 PHE_SOLTU	P34094 solanum tub
30	82	5.5	229	1 RPE_CHLPN	092829 chlamydia p
31	82	5.5	492	1 6PGD_SCHPO	P78812 schizosacch
32	82	5.5	540	1 CH60_THERB	060024 thermoaer
33	81.5	5.5	636	1 GYRB_THDMA	P77993 thermotoga

34	81.5	5.5	1002	1 YEMA_DROME	P25992 drosophila
35	81	5.4	309	1 CC2B_ARATH	P25859 arabidopsis
36	81	5.4	459	1 Y819_PYRHO	058549 pyrococcus
37	81	5.4	560	1 VAOX_PENSI	P56216 penicillium
38	81	5.4	864	1 CHEA_BORBU	044737 borrelia bu
39	81	5.4	885	1 PMC2_HUMAN	001780 homo sapien
40	80.5	5.4	399	1 RRP2_HUMAN	P22044 human parat
41	80.5	5.4	451	1 NFS1_MOUSE	092143 mus musculu
42	80.5	5.4	473	1 VL2_HPV28	P50799 human papil
43	80.5	5.4	543	1 PROA_LEGPN	P21347 legionella
44	80.5	5.4	733	1 RRP6_YEAST	012149 saccharomyc
45	80	5.4	574	1 DPO1_AQUAE	067779 aquifex aeo

ALIGNMENTS

RESULT 1
WRN_HUMAN STANDARD; PRT; 1432 AA.
AC Q14191;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Werner syndrome helicase.
GN WRN OR RECQL2 OR RECQ3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9618115; PubMed=8602509;
RA Yu C.-E., Oshima J., Fu Y.-H., Wijsman E.M., Hisama F., Alisch R.,
RA Matthews S., Nakura J., Miki T., Quais S., Martin G.M., Mulligan J.,
RA Scheißenberg G.D.;
RT "Positional cloning of the Werner's syndrome gene";
RL Science 272:258-262(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Pieper B.W., Gayle M., Brady W., Swartz A., Gillett L.A., Alisch R.S.,
RA Mulligan J., Galas D., Fu Y.-H.;
RT "Genomic structure of the human Werner's gene and cloning of the
RT mouse homolog";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SUBCELLULAR LOCATION.
RX MEDLINE=98284027; PubMed=9618508;
RA Marciniak R.A., Lombard D.B., Johnson F.B., Guarente L.;
RT "Nucleolar localization of the Werner syndrome protein in human
RT cells";
RL Proc. Natl. Acad. Sci. U.S.A. 95:6887-6892(1998).
RN [4]
RP REPEATS.
RX MEDLINE=99160561; PubMed=10049920;
RA Kusano K., Berres M.E., Engels W.R.;
RT "Evolution of the RECQ family of helicases: A drosophila homolog,
RT Dmblm, is similar to the human Bloom syndrome gene";
RL Genetics 151:1027-1039(1999).
RN [5]
RP REVIEW ON VARIANTS.
RX MEDLINE=99235545; PubMed=10220139;
RA Moser M.J., Oshima J., Monnat R.J., Jr.;
RT "WRN mutations in Werner syndrome";
RL Hum. Mutat. 13:271-279(1999).
RN [6]
RP VARIANT ARG-1367.
RX MEDLINE=9713161; PubMed=9021029;
RA Ye L., Miki T., Nakura J., Oshima J., Kamino K., Rakugi H.,
RA Ikegami H., Hiyaki J., Edland S.D., Martin G.M., Ogihara T.;
RT "Association of a polymorphic variant of the Werner helicase gene with
RT myocardial infarction in a Japanese population";
RL Am. J. Med. Genet. 68:494-498(1997).

RN [7]
 RP ERRATUM.
 RA Ye L., Miki T., Nakura J., Oshima J., Kamino K., Rakugi H.,
 RA Ikegami H., Higaki J., Egland S.D., Martin G.M., Ogihara T.,
 RL Am. J. Med. Genet. 70:103-103(1997).
 RN [8]
 RP VARIANTS ILE-387 AND LEU-1074.
 RX MEDLINE-98111850; PubMed-9450180;
 RA Medsitzler C., Rupplitsch W., Weirich-Schwaiger H., Weirich H.G.,
 RA Jabkowsky J., Klein G., Schweizer M., Hirsch-Kauffmann M.,
 RT "Werner syndrome: characterization of mutations in the WRN gene in an
 RT affected family.";
 RL Eur. J. Hum. Genet. 5:364-370(1997).
 RN [9]
 RP VARIANT ILE-387.
 RA Vidal V., Bay J.-O., Champomier F., Grancho M., Beauville L.,
 RA Glowaczower C., Lemery D., Ferreira M., Bignon Y.-J.,
 RT "The 1396del A mutation and a missense mutation of a rare polymorphism
 RT of the WRN gene detected in a French Werner family with a severe
 RT phenotype and a case of an unusual vulvar cancer.";
 RL Hum. Mutat. 11:413-414(1998).
 RN [10]
 RP VARIANTS ALA-324 AND ARG-1367.
 RX MEDLINE-99167244; PubMed-10069711;
 RA Casto E., Ogburn C.E., Hunt K.E., Tlilis R., Louhija J.,
 RA Penttinen R., Erkkola R., Panduro A., Riestra R., Plussan C.,
 RA Deeb S.S., Wang L., Edland S.D., Martin G.M., Oshima J.,
 RT "Polymorphisms at the Werner locus: I. Newly identified polymorphisms,
 RT ethnic variability of 1367C>G/Arg, and its stability in a population
 RT of Finnish centenarians.";
 RL Am. J. Med. Genet. 82:399-403(1999).
 CC -1- FUNCTION: ESSENTIAL FOR THE FORMATION OF DNA REPLICATION FOCAL
 CC CENTERS; STABLY ASSOCIATES WITH FOCI ELEMENTS GENERATING BINDING
 CC SITES FOR RP-A. EXHIBITS A MAGNESIUM-DEPENDENT ATP-DEPENDENT DNA-
 CC HELICASE ACTIVITY. MAY BE INVOLVED IN THE CONTROL OF GENOMIC
 CC STABILITY (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
 CC -1- DISEASE: DEFECTS IN WRN ARE THE CAUSE OF WERNER SYNDROME (WS): A
 CC RARE AUTOSOMAL RECESSIVE PROGEROID SYNDROME CHARACTERIZED BY THE
 CC PREMATURE ONSET OF MULTIPLE AGE-RELATED DISORDERS, INCLUDING
 CC ATHEROSCLEROSIS, CANCER, NON-INSULIN-DEPENDENT DIABETES MELLITUS
 CC (NIDDM), OCULAR CATARACTS AND OSTEOPOROSIS. THE MAJOR CAUSE OF
 CC DEATH (AT A MEDIAN AGE OF 47) IS MYOCARDIAL INFARCTION (MI).
 CC -1- SIMILARITY: BELONGS TO THE HELICASE FAMILY. RECO SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 HRDC DOMAIN.
 CC -1- DATABASE: NAME=WRN; NOTE=WRN mutation db (Werner disease):
 CC WWW="http://www.pathology.washington.edu/werner/ws_wrn.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: L76937; AAC41981.1; -
 DR EMBL: AF091214; AAC63361.1; -
 DR EMBL: AF181897; AAF06162.1; -
 DR EMBL: AF181896; AAF06162.1; JOINED.
 DR MIM: 604611; -
 DR MIM: 277700; -
 DR InterPro: IPR002562; 3_5_exonuclease.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR002121; HRDC.
 DR InterPro: IPR001650; Helicase_C.
 DR Pfam: PF01612; 3_5_exonuclease; 1.
 DR Pfam: PF00270; DEAD; 1.
 DR Pfam: PF00271; helicase_C; 1.
 DR Pfam: PF00570; HRDC; 1.
 DR SMART: SM00474; 35XOC; 1.
 DR SMART: SM00487; DEXDC; 1.
 DR SMART: SM00490; HELICc; 1.

DR SMART: SM00341; HRDC; 1.
 KW Hydrolyase; Helicase; ATP-binding; DNA-binding; Nuclear protein;
 KW Polymorphism; Repeat.
 FT DOMAIN 424 477
 FT
 FT
 FT REPEAT 424 450
 FT REPEAT 451 477
 FT DOMAIN 507 510
 FT DOMAIN 1150 1229
 FT NP_BIND 571 578
 FT SITE 668 671
 FT VARIANT 324 324
 FT
 FT VARIANT 387 387
 FT
 FT VARIANT 1074 1074
 FT
 FT VARIANT 1367 1367
 FT
 FT
 FT SEQUENCE 1432 AA; 162494 MM; DE02C0059F7B62EB CRC64;
 SQ
 Query Match 23.6%; Score 352; DB 1; Length 1432;
 Best Local Similarity 37.7%; Pred. No. 8.8e-22;
 Matches 75; Conservative 41; Mismatches 75; Indels 8; Gaps 3;
 QY 88 NPFAMRPGRIILSKATREKRAMOLIKVLDKRDSEGIAPVGLDIEMRSPRKGLPG 147
 DB 43 DLPLEFTGSIVSYDSDSFLSE-----DISMSISDGDVGVFDMENPPLYNRKL-G 95
 QY 148 KVATVQICVDNVCYDMHIFHSGI-POSTQHLIEDSTLVKVGIGIDGSKFLHDYGSVI 206
 DB 96 KVALIDQVSEKSCYLFHVSMSGVFPQGLKMLENNKAVKAGVIGDQWKLRLDPIKL 155
 QY 207 KQVEDLSLANQKIGGDKKGLASLETETLYCKELLPNRRLGNMEFFPLSKOLOYAAT 266
 DB 156 KNFVELTDVANKRLKCTEWSLSLVKHLKQLDLKDRCSIRCSNWSKFLTEQOKLYAAT 215
 QY 267 DAYASHLYLVKLDLPDAV 285
 DB 216 DAYAGFIYRNLEILDIV 234
 RESULT 2
 WRN_MOUSE STANDARD; PRT; 1401 AA.
 ID WRN_MOUSE 009053; 009050; 092242;
 AC 009053; 009050; 092242;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Werner syndrome helicase homolog.
 GN WRN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCB1_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BAIB/C; TISSUE=Testis, and Spleen;
 RX MEDLINE-97288537; PubMed-9143515;
 RA Imanura O., Ichikawa K., Yamabe Y., Goto M., Sugawara M.,
 RA Furuchi Y.,
 RT "Cloning of a mouse homologue of the human Werner syndrome gene and
 RT assignment to 8A4 by fluorescence in situ hybridization.";
 RL Genomics 41:298-300(1997).
 RN [2]
 RP SUBCELLULAR LOCATION.
 RX MEDLINE-98284027; PubMed-9618508;
 RA Marciniak R.A., Lombard D.B., Johnson F.B., Guarente L.,
 RT "Nucleolar localization of the Werner syndrome protein in human
 RT cells".
 RL Proc. Natl. Acad. Sci. U.S.A. 95:6887-6892(1998).


```

RN [3]
SEQUENCE FROM N.A.
RA Peepers B.W., Gayle M., Brady W., Swartz A., Gillett L.A., Alisch R.S.,
RA Mulligan J., Galas D., Fu Y.-H.;
RT "Genomic structure of the human Werner's gene and cloning of its mouse
RT homolog.";
CC Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAY BE INVOLVED IN THE CONTROL OF GENOMIC STABILITY.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE HELICASE FAMILY. RECD SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 1 HRDC DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: D86527; BAA20270.1; -
DR EMBL: D86526; BAA20269.1; -
DR EMBL: AF091215; AAC78077.1; -
DR MGD: MGI:109635; Wtn
DR InterPro: IPR002562; 3_5_exonuclease.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR001211; HRDC.
DR InterPro: IPR001650; Helicase_C.
DR Pfam: PF01612; 3_5_exonuclease; 1.
DR Pfam: PF00270; DEAD; 1.
DR Pfam: PF00271; Helicase_C; 1.
DR Pfam: PF00370; HRDC; 1.
DR SMART: SM00474; 35EXOC; 1.
DR SMART: SM00487; DEXDC; 1.
DR SMART: SM00490; HELIC; 1.
DR SMART: SM00341; HRDC; 1.
KM Helicase; ATP-binding; Nuclear protein.
FT NP_BIND 535 542 ATP (BT SIMILARITY).
FT SITE 632 635 DEAD BOX.
FT DOMAIN 1115 1194 POLY-SER.
FT DOMAIN 1387 1390 HRDC.
FT CONFLICT 101 101 N -> S (IN REF. 3).
FT CONFLICT 228 228 V -> A (IN REF. 3).
FT CONFLICT 250 250 L -> S (IN REF. 3).
FT CONFLICT 452 452 M -> V (IN REF. 3).
FT CONFLICT 459 459 K -> T (IN REF. 3).
FT CONFLICT 468 468 C -> R (IN REF. 3).
FT CONFLICT 619 619 K -> Q (IN REF. 3).
FT CONFLICT 800 800 Q -> K (IN REF. 3).
FT CONFLICT 1021 1021 L -> S (IN REF. 3).
FT CONFLICT 1145 1145 A -> T (IN REF. 3).
FT CONFLICT 1181 1182 VG -> LE (IN REF. 3).
FT CONFLICT 1252 1252 V -> A (IN REF. 3).
FT CONFLICT 1308 1308 I -> L (IN REF. 3).
FT CONFLICT 1356 1356 V -> A (IN REF. 3).
SQ SEQUENCE 1401 AA; 157256 MW; 94906092467FB8C CRC64;

Query Match 22.7%; Score 339; DB 1; Length 1401;
Best Local Similarity 36.6%; Pred. No. 1.1e-20;
Matches 74; Conservative 42; Mismatches 72; Indels 14; Gaps 5;

OY 88 NPAAMRGGLILYKATATVDKRAMOLIKVLDTRKDESGIAFVGLDIEMRPSFKGVLP 147
DB 37 NLFLEPPGGIIVSYEASD-----CSFLSEDISMRLSDGV--VGFDMEPPIYK----PG 86
OY 148 K---VATVOICVDSNCDVNHFIHSGI-POSLOHLEDSTLVKVGITIDDSVKVLFHDYG 203
DB 87 KRRRAVAIOLCVSEENKCYLRHISMSVFPGLKMLENKSIRKAGVIEEDQWKLRLDPD 146
OY 204 VSIRKDEEDSLDANOKIGDKRWGLASLTFTLVCKELKFNRIKIGWMEPYPLSKOOLQY 263
DB 147 VKLESFVELTDVANERKKAEMTWSLNGVLKHYGLKOLKDKSIRCSWMSNPLTEDKLY 206
```

```

OY 264 AATAYASWHLIKYKVLDPAY 285
DB 207 AATDAYAGLIYKRLGNIGDTV 228

RESULT 3
Y068_CAEEL STANDARD; PRT; 910 AA.
AC P34607;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Hypothetical 105.6 kDa protein ZK1098.8 in chromosome III.
GN ZK1098.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=94150718; Pubmed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kersey J., Kirsten J., Laister N.,
RA Lettelle P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA Sims M., Smaildon N., Smith A., Smith M., Sonnhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Wellstock L., Wilkinson-Sproat J.,
RA Wooldam P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Mature 368,32-38(1994).
CC -!- SIMILARITY: TO RIBONUCLEASE D.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: Z22176; CAA80137.1; -
DR PTR: S40930; S40930.
DR WormPep: ZK1098.8; CE00370.
DR InterPro: IPR002562; 3_5_exonuclease.
DR Pfam: PF01612; 3_5_exonuclease; 1.
DR SMART: SM00474; 35EXOC; 1.
KM Hypothetical protein.
FT NP_BIND 910 AA; 105569 MW; 5512D15423517FCD CRC64;

Query Match 9.2%; Score 136.5; DB 1; Length 910;
Best Local Similarity 24.8%; Pred. No. 0.00094;
Matches 77; Conservative 34; Mismatches 110; Indels 89; Gaps 13;

OY 12 TEEFLAIDAIEASYNRSRSSSSSSAAPTVOATTSVHGHEEDPNQIPNNIRQLPSIT 71
DB 333 TEKO--IDAICFAIYLGIASSSS-----SKTATVDKRAMOLIKVLD 119
OY 72 SSTSYKRPFLSRCAFRNFPAMRFEGRLTY-----SKTATVDKRAMOLIKVLD 119
DB 369 SKLREAKELLVRRKTLQV---LNGEOLVFNENRRQTQIMWKTESMN-----YLCEIK 421
OY 120 TKRDSEGIAFVGLDIEMRPSFKGVLPKQVATVOI-----CVD---SNYCDVMI 166
DB 422 SLSDPAPVYVGFDSWKPSNLTFAVHDSKIALIQLFEKNCVWLVDCELEKANMAD---- 477
OY 167 FHSIGTPOSLOHLEDSTLVKVGITIDG-----SVKLFHDYGVSIKQVDELSD 214
```

Db 478 --DMMOKFASRLFGDSFVKVYGFDMRNDLDMATIPALKSGSMKI-----EDTKNAFDLKR 530
QY 215 LANOKIGD-----KKWGLASLFTLYCKELTKPNRIKLGWEPFLSKOOLYOAT 266
Db 531 LAENVCDIDMEILELPKRTKRLADTHYILGLEDKTE--QCSNMOCRLPKKQIYVAAL 588
QY 267 DAVASMHLYK 276
Db 589 DAVVVFETFK 598

RESULT 4
CPVL_RAT
ID CPVL_RAT STANDARD; PRT; 508 AA.
AC P22443;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Cytochrome P450 19 (Aromatase) (EC 1.14.14.1) (CYPXIX) (Estrogen synthetase) (P-450AROM).
GN CYP19 OR AROM.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=90220647; PubMed=2157976;
RA Hickey G.J., Krasnow J.S., Beattie W.G., Richards J.S.;
RT "Aromatase cytochrome P450 in rat ovarian granulosa cells before and after luteinization: adenosine 3',5'-monophosphate-dependent and independent regulation. Cloning and sequencing of rat aromatase cDNA and 5' genomic DNA."
RT MOL. Endocrinol. 4:3-12(1990).
CC -1- FUNCTION: CATALYZES THE FORMATION OF AROMATIC C18 ESTROGENS FROM C19 ANDROGENS.
CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH + oxidized flavoprotein + H(2)O.
CC -1- SUBCELLULAR LOCATION: Membrane-bound.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sid.ch](mailto:license@isb-sib.ch)).
CC -----
DR EMBL: M33986; AAA41044.1; -
DR PIR: A36121; A36121.
DR InterPro: IPR001128; Cyt_P450.
DR Pfam: PF00067; P450. 1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; 1.
KW Electron transport; Oxidoreductase; Monooxygenase; Membrane; Heme.
FT BINDING 437 437 HEME.
SQ SEQUENCE 508 AA; 58411 MW; C0ADF0F0D80AB352 CRC64;

Query Match 6.4%; Score 95; DB 1; Length 508;
Best Local Similarity 23.1%; Pred. NO. 1.4;
Matches 42; Conservative 25; Mismatches 41; Indels 74; Gaps 9;

QY 93 REGGRILYKSTATEVDKRAQLLVLDTRKDESGIAFVGIDIEW---RPFKRGVL-PGR 148
Db 115 RFG-----SKRGLQCIQM---HENGILFNPNPSLMRYVPFPMKALGPG 157
QY 149 VATVQICVD-----SNQDVNHLFHSQIPQSLQHLIED-STLVKVGIGID 192
Db 158 IRNVEVCVESIKQHLDRLDGVTNDSGIVDVYTL-----MRHIMIDTSNTLFLGIPLD 209

QY 193 GDSYK-----LFHDYGSINDVED---LSDLANOKIGD 223
Db 210 ESSIVKKIQQYFNAMQALLIKPNIFKISMLYKRYEVSVDLDEILEYKRRQYSSA 269
QY 224 K 225
Db 270 EK 271

RESULT 5
POLG_STEVN
ID POLG_STEVN STANDARD; PRT; 1525 AA.
AC P09732; Q88781; Q88782; Q88783; Q88784; Q88785; Q88786; Q88787;
AC Q88788;
DT 01-MAR-1989 (Rel. 10, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polyprotein [contains: capsid protein C (Core protein); Matrix protein (Envelope protein M); Major envelope protein E; Nonstructural proteins NS1, NS2A, NS2B; Protease/helicase (EC 3.4.21.98) (NS3)] (Fragment).
DE St. Louis encephalitis virus (strain M51-7).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; OC Flavivirus.
OX NCBI_TaxID=11081;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87121217; PubMed=3027980;
RA Trent D.W., Kinney R.M., Johnson B.J.B., Vorndam A.V., Grant J.A., Daubel V., Rice C.M., Hahn C.;
RT "Partial nucleotide sequence of St. Louis encephalitis virus RNA: structural proteins, NS1, NS2A, and NS2B."
RT Virology 156:293-304(1987).
CC -1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
CC NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF FOUR PEPTIDE BONDS IN THE VIRAL PRECURSOR POLYPROTEIN, COMMONLY WITH ASP OR GLU IN THE P6 POSITION, CYS OR THR IN P1 AND SER OR ALA IN P1'.
CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND RNA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M16614; AAA4786.1; -
DR PIR: A27531; GNMV55.
DR HSSP: P14336; ISVB.
DR InterPro: IPR000069; FLAVI_M.
DR InterPro: IPR001157; FLAVI_NS1.
DR InterPro: IPR000752; FLAVI_NS2A.
DR InterPro: IPR000487; FLAVI_NS2B.
DR InterPro: IPR001122; FLAVI_capsid.
DR InterPro: IPR000336; FLAVI_capsid.
DR InterPro: IPR002535; FLAVI_propep.
DR Pfam: PF01003; FLAVI_capsid; 1.
DR Pfam: PF00869; FLAVI_glycoprot; 1.
DR Pfam: PF01004; FLAVI_M; 1.
DR Pfam: PF01005; FLAVI_NS2A; 1.
DR Pfam: PF01002; FLAVI_NS2B; 1.
DR Pfam: PF01570; FLAVI_propep; 1.
DR ProDom: PD001496; FLAVI_NS1; 1.
DR ProDom: PD001556; FLAVI_glycoprote; 1.
KW Polyprotein; Glycoprotein; Core protein; Coat protein; Envelope protein; Hydroxylase; Helicase; ATP-binding; Transmembrane; Nonstructural protein.

FT	INT_MET	1	1	REMOVED FROM CAPSID PROTEIN C BY THE
FT	CHAIN	1	121	CELLULAR AMINOPEPTIDASE.
FT	PROPEP	122	213	CAPSID PROTEIN C.
FT	CHAIN	214	288	ENVELOPE GLYCOPROTEIN M.
FT	CHAIN	289	789	MAJOR ENVELOPE PROTEIN E.
FT	CHAIN	790	1203	NONSTRUCTURAL PROTEIN NS1.
FT	CHAIN	1204	1368	NONSTRUCTURAL PROTEIN NS2A.
FT	CHAIN	1369	1499	NONSTRUCTURAL PROTEIN NS2B.
FT	CHAIN	1500	>1525	PROTEASE/HELICASE (NS3).
FT	TRANSMEM	108	119	POTENTIAL.
FT	TRANSMEM	253	268	POTENTIAL.
FT	TRANSMEM	274	288	POTENTIAL.
FT	TRANSMEM	751	762	POTENTIAL.
FT	TRANSMEM	768	787	POTENTIAL.
FT	DISULFID	1173	1188	POTENTIAL.
FT	DISULFID	291	318	BY SIMILARITY.
FT	DISULFID	348	404	BY SIMILARITY.
FT	DISULFID	362	393	BY SIMILARITY.
FT	DISULFID	380	409	BY SIMILARITY.
FT	DISULFID	478	576	BY SIMILARITY.
FT	DISULFID	593	624	BY SIMILARITY.
FT	CARBOHYD	136	136	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	269	269	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	442	442	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	919	919	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	964	964	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	996	996	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1189	1189	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	NON_TER	1525	1525	N-LINKED (GLCNAC. . .) (POTENTIAL).
SO	SEQUENCE	1525 AA;	167891 MW;	E1A373F1E511159F CRC64;
Query Match		6.3%;	Score 93.5;	DB 1; Length 1525;
Best Local Similarity		22.3%;	Pred. NO. 8.4;	
Matches		47; Conservative	34; Mismatches	85; Indels 45; Gaps 11;
Db	82 SRCARNPPARFGRILXSKATEVDRKAMOLIKVLDTRKDE---	SGIAFGGLDIE-WR	137	
Db	764 ARDRISITLLVAGSILFLFVATSVQADSGC-----AISLQREELKCGGIFVYNDVEKKK	818		
Qy	138 PSFRKGVV--PGKAVTQICVSNDCDV4-----HIFHSIGPOSLOHLEIDSTLVKCVI	189		
Db	819 SDYKFFPLPPGLAVIVDEAHANGYCGIRSTSRLEHLMENQRELNALFEDNE-IDLSV	877		
Qy	190 GIDGDSVKLFHDYGVSIKDEVDLSLANOKIGDKKMGASLTETLVCKELLPNRIRLG	249		
Db	878 VVOEDP--KYKRAPRRLKKLEDELDP-----GKKMKGTLEVP-----RLG	918		
Qy	250 NMEFY---PLSNQQLQVYATDAYVASMHLYKV	277		
Db	919 NMTFVVDGPETKE-----CPTANRRAMNSFKV	944		
RESULT 6				
FA5_HUMAN	STANDARD:	PRT:	2224 AA.	
AC	P12259: 014285;			
DT	01-OCT-1989 (Rel. 12, Created)			
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Coagulation factor V precursor (Activated protein C cofactor).			
FN	F5.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RN	SEQUENCE FROM N.A.			
RX	MEDLINE=92232668; PubMed=1567832;			
RT	Cripe L.D., Moore K.D., Kane W.H.;			
RT	"Structure of the gene for human coagulation factor V.";			
RT	Biochemistry 31:3777-3785(1992).			
LN	[2]			

CC -1- DISEASE: OMREN PARAHOMOPHILIA, AN HEMORRHAGIC DIATHESIS, IS DUE
CC TO A DEFICIENCY OF FACTOR V. OTHER DEFECTS IN F5 RESULTS IN A
CC FORM OF THROMBOPHILIA KNOWN AS APC RESISTANCE (APCR). THE APCR
CC MUTATION IS FOUND IN ABOUT 5% OF THE POPULATION WHICH SUGGEST THAT
CC A SLIGHT THROMBOTIC TENDENCY MAY CONFER SOME ADVANTAGE IN FETAL
CC IMPLANTATION.
CC -1- SIMILARITY: CONTAINS 3 F5/8 TYPE A DOMAINS; EACH IS COMPOSED OF
CC 2 PLASTOCYANIN-LIKE REPEATS.
CC -1- SIMILARITY: STRONG, TO COAGULATION FACTOR VIII.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L32779; AAB59401.1; -
DR EMBL: L32755; AAB59401.1; JOINED.
DR EMBL: L32756; AAB59401.1; JOINED.
DR EMBL: L32757; AAB59401.1; JOINED.
DR EMBL: L32758; AAB59401.1; JOINED.
DR EMBL: L32759; AAB59401.1; JOINED.
DR EMBL: L32760; AAB59401.1; JOINED.
DR EMBL: L32761; AAB59401.1; JOINED.
DR EMBL: L32762; AAB59401.1; JOINED.
DR EMBL: L32763; AAB59401.1; JOINED.
DR EMBL: L32764; AAB59401.1; JOINED.
DR EMBL: L32765; AAB59401.1; JOINED.
DR EMBL: L32766; AAB59401.1; JOINED.
DR EMBL: L32767; AAB59401.1; JOINED.
DR EMBL: L32768; AAB59401.1; JOINED.
DR EMBL: L32769; AAB59401.1; JOINED.
DR EMBL: L32770; AAB59401.1; JOINED.
DR EMBL: L32771; AAB59401.1; JOINED.
DR EMBL: L32772; AAB59401.1; JOINED.
DR EMBL: L32773; AAB59401.1; JOINED.
DR EMBL: L32774; AAB59401.1; JOINED.
DR EMBL: L32775; AAB59401.1; JOINED.
DR EMBL: L32776; AAB59401.1; JOINED.
DR EMBL: L32777; AAB59401.1; JOINED.
DR EMBL: L32778; AAB59401.1; JOINED.
DR EMBL: M16967; AAB52424.1; -
DR EMBL: M14355; AAB59532.1; -
DR PIR: A25897; A25897.
DR PIR: A28028; A28028.
DR PDB: 1CZS; 26-NOV-99.
DR PDB: 1CZT; 26-NOV-99.
DR PDB: 1CZV; 26-NOV-99.
DR MIM: 134400; -
DR MIM: 188055; -
DR MIM: 227310; -
DR MIM: 227400; -
DR InterPro: IPR001117; Cu-oxidase.
DR InterPro: IPR000421; FA58_C.
DR Pfam: PFO0394; Cu-oxidase; 3.
DR Pfam: PFO0754; F5_F8_type_C; 2.
DR SMART: SM00231; FA58C_2.
DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 2.
DR PROSITE: PS01285; FA58C_1; 2.
DR PROSITE: PS01285; FA58C_2; 2.
KW Blood coagulation; Glycoprotein; Sulfation; Calcium; Signal; Zymogen;
KW Repeat; Polymorphism; Glycoprotein; Thrombophilia; 3D-structure.
FT SIGNAL 1 28
FT CHAIN 29 2224
FT CHAIN 29 737
FT PEPTIDE 738 1573
FT CHAIN 1574 2224
FT DOMAIN 30 329
FT DOMAIN 30 193
FT DOMAIN 203 329
PLASTOCYANIN-LIKE 1.
PLASTOCYANIN-LIKE 2.

FT DOMAIN 348 684 F5/8 TYPE A 2.
FT DOMAIN 348 526 PLASTOCYANIN-LIKE 3.
FT DOMAIN 536 684 PLASTOCYANIN-LIKE 4.
FT DOMAIN 692 1573 B.
FT DOMAIN 895 928 2 x 17 AA TANDEM REPEATS.
FT REPEAT 895 911 1-1.
FT REPEAT 912 928 1-2.
FT SIMILAR 1135 1148 TO 14 AA REPEATS IN BOVINE FA5.
FT DOMAIN 1185 1501 35 x 9 AA APPROXIMATE TANDEM REPEATS OF
[TNP]-L-S-P-D-L-S-Q-T.
FT REPEAT 1185 1193 2-1.
FT REPEAT 1194 1202 2-2.
FT REPEAT 1203 1211 2-3.
FT REPEAT 1212 1220 2-4.
FT REPEAT 1221 1229 2-5.
FT REPEAT 1230 1238 2-6.
FT REPEAT 1239 1247 2-7.
FT REPEAT 1248 1256 2-8.
FT REPEAT 1257 1265 2-9.
FT REPEAT 1266 1274 2-10.
FT REPEAT 1275 1283 2-11.
FT REPEAT 1284 1292 2-12.
FT REPEAT 1293 1301 2-13.
FT REPEAT 1302 1310 2-14.
FT REPEAT 1311 1319 2-15.
FT REPEAT 1320 1328 2-16.
FT REPEAT 1329 1337 2-17.
FT REPEAT 1338 1346 2-18.
FT REPEAT 1347 1355 2-19.
FT REPEAT 1356 1364 2-20.
FT REPEAT 1365 1373 2-21.
FT REPEAT 1374 1382 2-22.
FT REPEAT 1383 1391 2-23.
FT REPEAT 1392 1400 2-24.
FT REPEAT 1401 1409 2-25.
FT REPEAT 1410 1418 2-26.

Query Match 6.2%; Score 93; DB 1; Length 2224;
Best Local Similarity 19.2%; Pred. No. 15;
Matches 58; Conservative 42; Mismatches 124; Indels 78; Gaps 11;

QY 4 SNNIDAFTEBELLAIDAIIEASYNFESSSSSSSSAAATVQATSYVHGHEDDPNOIPNNIR 63
DB 126 ASYLDHTEPAEKM--DDAVAGREYETWEISIEDSGPT-----HDDP--CLTHIY 172
QY 64 RQLPRSTSTSYKRPFLSCRRANFPAMFRGRILYSKTATEYDKRAMOLIKVLDTRKD 123
DB 173 YSHENLIEDNLSGLIGLLICK-----GLTREGTQKTFDKQIVLLFAVPDESK- 222
QY 124 ESGIAFVGLDIEWRPSFR-----KGVLPKVAATVQICVDSNYCDVMHIFHSGIPQSLQHL 178
DB 223 -----SWSQSSSLMYTVNGVNGTMDITVCA-----HDSIMWHLGM 260
QY 179 IEDSTLVKVGIGIDGDSVKLPFDYGVSIKDYEDLSLANKIGDKKMGSLAST----- 232
DB 261 SSGPEL--FSIHFGQVLEQNNHKVSAITLVSAITSTANNTVQPEGKWIISLTPPHLQA 318
QY 233 -----ETLVCKELKPNRIKGMFEPYLSKOOL-QVAAAT-----DAVASW 272
DB 319 GMAQYIDIKRCPKTRKLKNTREORHMKRWEYFLAAEYVINDYAPVPIANNDKRYRSO 378
QY 273 HL 274
DB 379 HL 380

RESULT 7
ID IRA2_YEAST STANDARD; PRT; 3079 AA.
AC P19158; O08239;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)

```

DE      Inhibitory regulator protein IRA2.
DN      IRA2 OR GLC4 OR CCS1 OR YOL081W OR 000985.
OS      Saccharomyces cerevisiae (Baker's yeast).
OC      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OX      Saccharomycetales; Saccharomycetaceae; Saccharomyces.
ON      NCBI_TaxID=4932;
OR      [1]
RX      MEDLINE=90318397; PubMed=2164637;
RA      Tanaka K., Nakafuku M., Tamanoi F., Kaziro Y., Matsumoto K., Toh-E A.;
RT      "IRA2, a second gene of Saccharomyces cerevisiae that encodes a
RT      protein with a domain homologous to mammalian ras GTPase-activating
RT      protein."
RL      Mol. Cell. Biol. 10:4303-4313(1990).
RN      [2].
RP      SEQUENCE OF 1-2423 FROM N.A.
RC      STRAIN-S288C / F11679;
RA      MEDLINE=95208358; PubMed=7900427;
RN      Zimstein E., Griffin H., Schwizler M.;
RT      "Sequence of a 10.27 kb segment on the left arm of chromosome XV from
RT      Saccharomyces cerevisiae includes part of the IRA2 gene and a
RT      putative new gene."
RL      Yeast 10:1383-1387(1994).
RN      [3]
RP      SEQUENCE OF 1982-3079 FROM N.A.
RA      MEDLINE=97321807; PubMed=9178509;
RN      Tzeremia M., Katsoulou C., Alexandraki D.;
RT      "Sequence analysis of a 33.2 kb segment from the left arm of yeast
RT      chromosome XV reveals eight known genes and ten new open reading
RT      frames including homologues of ABC transporters, inositol
RT      phosphatases and human expressed sequence tags."
RL      Yeast 13:583-589(1997).
RN      [4]
RP      IDENTIFICATION OF CCS1 AS IRA2.
RA      Bussereau F., Dupont C.H., Boy-Marcotte E., Maillet L., Jaquet M.;
RT      "The CCS1 gene from Saccharomyces cerevisiae which is involved in
RT      mitochondrial functions is identified as IRA2 an attenuator of RAS1
RT      and RAS2 gene products."
RL      Curr. Genet. 21:325-329(1992).
CC      -I- FUNCTION: INHIBITORY REGULATOR OF THE RAS-CYCCLIC AMP PATHWAY.
CC      -I- STIMULATES THE GTPASE ACTIVITY OF RAS PROTEINS.
CC      -I- SIMILARITY: CONTAINS 1 RAS-GAP DOMAIN.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M33779; AAA34710.1; -
DR      EMBL; X83121; CAA58201.1; -
DR      EMBL; X75449; CAA53202.1; -
DR      EMBL; Z74823; CAA99093.1; -
DR      PIR; S11190; RGEYI2.
DR      SGD; S0005441; IRA2.
DR      InterPro; IPR001936; RasGAP.
DR      Pfam; PF00616; RasGAP; 1.
DR      SMART; SM00323; RasGAP; 1.
DR      PROSITE; PS00509; RAS_GTPASE_ACTIV_1; 1.
DR      PROSITE; PS001018; RAS_GTPASE_ACTIV_2; 1.
KW      GTPase activation.
FT      DOMAIN 1701 1890 RAS-GAP.
FT      DOMAIN 399 409 POLY-SER.
FT      DOMAIN 412 416 POLY-SER.
FT      DOMAIN 520 528 POLY-ALA.
FT      DOMAIN 2469 2472 POLY-LEU.
FT      CONFLICT 2317 2317 I -> K (JIN REF. 3).
SO      SEQUENCE 3079 AA; 351631 MW; 651EB2A2EEB479C0 CRC64;

```

Best Local Similarity 19.3%: Pred. No. 24:
Matches 57; Conservative 43; Mismatches 112; Indels 84; Gaps 10.

RESULT	8
LEU2_SCHPO	
ID	LEU2_SCHPO
AC	014289;
DT	15-JUL-1998 (Rel. 36, Created)
DR	15-JUL-1998 (Rel. 36, Last sequence update)
DW	15-JUL-1998 (Rel. 36, Last annotation update)
DE	3-ISOPROPYLMALATE dehydratase (EC 4.2.1.33) (Isopropylmalate isomerase) (Alpha-IPM isomerase) (IPMI).
GN	SPAC9E9.03.
OS	Schizosaccharomyces pombe (Fission yeast).
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
CC	Schizosaccharomycetales; Schizosaccharomycetaceae;
OX	Schizosaccharomycetes.
RN	NCB1_TaxID=4896;
RP	[1]
RC	SEQUENCE FROM N.A.
RL	STRAIN=972;
RA	Mcdougall R., Barrell B.G., Rajandream M.A., Wood V.;
CC	Submitted (SEP-1997) to the EMBL/Genbank/DDBJ databases.
CC	-I CATALYTIC ACTIVITY: 3-ISOPROPYLMALATE = 2-ISOPROPYLMALATE + H(2)O (ALSO CATALYZES 2-ISOPROPYLMALATE + H(2)O = 3-HYDROXY-4-METHYL-3-CARBOXPENTANONE).
CC	-I PATHWAY: SECOND STEP IN LECICINE BIOSYNTHESIS.
CC	-I SIMILARITY: BELONGS TO THE ACONITASE/IPM ISOMERASE FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright.. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce,
CC	or send an email to license@isb-sib.ch).
CC	-----
CC	EMBL; Z99262; CAB16402.1; -
DR	InterPro: IPR001030; Aconitase.
DR	InterPro: IPR000573; Aconitase_C.
DR	Pfam: PF00330; aconitase; 1.
DR	Pfam: PF00694; Aconitase_G; 1.
DR	PRINTS: PR00415; ACONITASE.
DR	PRODOM: PD000511; Aconitase; 1.
DR	PROSITE: PS00450; ACONITASE_1; 1.
DR	PROSITE: PS01244; ACONITASE_2; 1.
KW	Lecicine biosynthesis; Lyase;
FT	MetAl 359 359 Iron-Sulfur; 4Fe-4S.
ET	MetAl 420 420 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
ET	MetAl 423 423 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
SO	SEQUENCE 758 AA; 82782 MW; CE78C36828380EA7 CR664; (BY SIMILARITY).

Query Match 6.2%; Score 92; DB 1; Length 758;
 Best Local Similarity 18.2%; Pred. No. 4.4;
 Matches 61; Conservative 53; Mismatches 107; Indels 114; Gaps 17;

QY 34 SSSSAAPTQATTS-----VHGHEDDPNOIPNNIR-ROLPRS-----ITSSTSYKREPL 81
 DB 292 NADAVLPTVWGTSPDVIPIINGNIPDPAHVKNVRAASTORSLEYGLKPNSTIVSYPI 351
 QY 82 SR-----CRARNFPAMRFGGRILYKSTKATEVDRAM-----OLIKYDTRDESGL--AF 129
 DB 352 DKAFITSCNTSRLEDLRLAAVYKGRVAAVNDAMIVPGSLVKW--AAEAGLDQIF 408
 QY 130 VGDIEMRP-----SFRKGVLPKGVATVOIC-----V 156
 DB 409 IEAGFDMREAGCSMCLGMNDQLKPYERCASSTNRNPEGRQAGKGRTHLYSPAMAAAI 468
 QY 157 DSNYCDVMHIF---HSGIPSLQHLIEDSTLVVVGIGIDGDSYKLRHDYGVSTKDEVL- 212
 DB 469 KGHLCNVREFGVDVSNQSPSIITNKNDPS-----HDVEG-----DIDLSDVDTADAV 516
 QY 213 --SDLANOKIGDKNKGLASL-----TETLVCKELKLP-NRIRLGN 250
 DB 517 TDADGATNAGSVSSGSACIPKFTVEGIAAPLPMAVNTDXTIIPKQFLTKTKTGLGQ 576

QY 251 WEFPYPLSKOQLOVAAVDAYASWHLKYVKLDLPDPAV 285
 DB 577 FAFY-----EIRYDADG-----KEIPDFV 595

RESULT 9
 ID_VL2_HPV03 STANDARD; PRT; 473 AA.
 AC P36744;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Minor capsid protein L2.
 GN L2.
 OS Human papillomavirus type 3.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 RN NCBI_TaxID=10614;
 RP [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94265501; PubMed=8205838;
 RA Delius H., Hofmann B.;
 RT "Primer-directed sequencing of human papillomavirus types."
 RT Curr. Top. Microbiol. Immunol. 186:13-31(1994).
 CC CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: X74462; CAA52473.1; -
 DR PIR: S36553;
 DR InterPro: IPR000784; Late_L2.
 DR Pfam: PF00513; Late_protein_L2; 1.
 KW Coat protein; Late protein.
 SQ SEQUENCE 473 AA; 50662 MW; DD8298FE517A04DD CRC64;

Query Match 6.1%; Score 91.5; DB 1; Length 473;
 Best Local Similarity 21.6%; Pred. No. 2.6;
 Matches 66; Conservative 34; Mismatches 95; Indels 111; Gaps 14;

QY 1 MSSSSNIDDAFTEELALDAIEASYNFSRSSSSSAAPTQATTSVHGHEDDPNOIPN 60
 DB 153 VSTSTNSNPAFTPELSLEVV-----QNEVSGHILISFTSGTHYEE----- 195
 QY 61 NTRQLPRSTSTSYKRPPLSRARNPAM-RFGGRILYKSTKATEVDRAMQLIKVLD 119

DB 196 -----IPMEFASPGTETETIS---STPPVGVERIRAGPRILYSAVAYQY-----KVTD 239
 QY 120 ---TKRDESGIAFVGLDIEMRPS---FRKGLPGKAVTVOICVDSNYCDVMHIF----- 167
 DB 240 PAFLTRPRSLMTDPNPFEPEDETIIIFERYSPSOVP-----DSDFLDLRLRHAPALTS 293
 QY 168 -----HSGIPSLQHLIEDSTLVKVGIGIDGDSYKLFHDYGVSIKDVEDLSLANOKIG 221
 DB 294 RRGTVKTSRYGQKL-----SMRTSGKL-----GARVHYDLSPIGP----- 332
 QY 222 GDRKWKGLASLTETLVCKELKLPNRIKLNMEFPYPLSKOQLOVATDAYASWHLKYVKLD 281
 DB 333 -----TEDIEMEPILAP-----ASASAYDS--LYDYVADV 360

QY 282 PDVAVG 287
 DB 361 DDADIG 366

RESULT 10
 ID_CPVL_MOUSE STANDARD; PRT; 503 AA.
 AC P28649;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Cytochrome P450 19 (Aromatase) (EC 1.14.14.1) (CYP19X) (Estrogen
 DE synthetase) (P-450AROM).
 GN CYP19 OR AROM.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.
 RN NCBI_TaxID=10090;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Ovary;
 RX MEDLINE=91378248; PubMed=1897929;
 RA Terashima M., Toda K., Kawamoto T., Kuribayashi I., Ogawa Y.,
 RA Maeda T., Shizuta Y.;
 RT "Isolation of a full-length cDNA encoding mouse aromatase P450."
 RT Arch. Biochem. Biophys. 285:231-237(1991).
 CC CC
 CC -1- FUNCTION: CATALYZES THE FORMATION OF AROMATIC C18 ESTROGENS FROM
 CC C19 ANDROGENS.
 CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
 CC oxidized flavoprotein + H(2)O.
 CC -1- SUBCELLULAR LOCATION: Membrane-bound.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: D00659; BAA00551.1; -
 DR PIR: S13912; S13912.
 DR MGD: MGI:88587; Cyp19.
 DR InterPro: IPR001128; Cyt_P450.
 DR Pfam: PF00067; P450; 1.
 DR PRINTS: PR00385; P450.
 DR PROSITE: PS00086; CYTOCHROME_P450; 1.
 KW Electron transport; Oxidoreductase; Monooxygenase; Membrane;
 KW Heme.
 FT BINDING 437 437 HEME (BY SIMILARITY).
 SQ SEQUENCE 503 AA; 58015 MW; 737400389D1AEFFI CRC64;

Query Match 6.1%; Score 91; DB 1; Length 503;
 Best Local Similarity 23.2%; Pred. No. 3.1;
 Matches 38; Conservative 23; Mismatches 33; Indels 70; Gaps 8;

```

OY 93 REGGRILYKSTATEDVKNRMQLIKVLDTRKDESGIAFVGIDIEW---RPSFRKGV-L-PCK 148
DB 115 REF-----SKRGLOQICG-----HENGIIFFNNPNSLMRPIRPFMKALGPGL 157
OY 149 VAVVOICVD-----SNVCDVMHIFHSGIPQSLQHLIED-STLVKVGIGID 192
DB 158 VRAVEVCESIKOHLDRLEGEVDTSGYVDVLT-----MHIMIDTNSMLFLGIPLD 209

OY 193 GDSVK-----LFHDYGSVKIDVED 211
DB 210 ESAIVKIKGIGYFNAMQALLIKPNIFFKISWLYKRYERSVYDLMD 253

RESULT 11
DPOL_HAEIN STANDARD; PRT; 930 AA.
AC P43741;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA polymerase I (EC 2.7.7.7) (POL I).
GN POLA OR HI0856.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus
OC NCBI_TaxId=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RA MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uiterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghegan N.S.M.,
RA Guelm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RA "Whole-genome random sequencing and assembly of Haemophilus
RA influenzae Rd.";
RT Science 269:496-512(1995).
RL
CC -1- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE
CC EXHIBITS 3' TO 5' AND 5' TO 3' EXONUCLEASE ACTIVITY (BY
CC SIMILARITY).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate -> N diphosphate
CC + (DNA)(N).
CC -1- SUBUNIT: SINGLE-CHAIN MONOMER WITH MULTIPLE FUNCTIONS.
CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-A FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL, U32767; AAC22515.1; -.
CC HSSP: P00582; IKFS.
CC TIGR: HI0856; -.
DR InterPro: IPR002562; 3-5-exonuclease.
DR InterPro: IPR002421; 5-3-exonuclease.
DR InterPro: IPR002298; DNA_POLI.
DR InterPro: IPR001098; DNA_POLI_A.
DR InterPro: IPR000513; Exo_N_I.
DR InterPro: IPR003584; HHH_2.
DR Pfam: PF01612; 3-5-exonuclease; 1.
DR Pfam: PF02136; 5-3-exonuclease; 1.
DR Pfam: PF02379; 5-3-exonuclease; 1.
DR Pfam: PF00476; DNA_POLI_A; 1.
DR PRINTS: PR00868; DNAPOLI.
DR SMART: SM00474; 35EXOC; 1.

```

```

DR SMART: SM00475; 53EXOC; 1.
DR SMART: SM00279; HHH2; 1.
DR SMART: SM00482; POLIAC; 1.
DR PROSITE: PS00447; DNA_POLYMERASE_A; 1.
KW Transferrase; DNA-directed DNA polymerase; DNA replication; DNA repair;
KW Hydrolyase; Exonuclease; DNA-binding; Complete proteome.
SQ SEQUENCE 930 AA; 103740 MW; 22654BB/CFF730B CRC64.

Query Match
Best local Similarity 20.7%; Score 89.5; DB 1; Length 930;
Matches 61; Conservative 42; Mismatches 93; Indels 99; Gaps 15;

OY 24 ASYNFSRSSSSSAAPTVOATTSVHGHEDPNOIPNNIRQLPRSTTSSTKRPPLSR 83
DB 280 ARYEFRKWLNEVMNGADSIQT-----EOP-----YKKNQYKATSODQ 318
OY 84 CRARNFPAMREGGRILYKSTATEDV-----KRAMOLIKVLDTRKDE-----SGIA 128
DB 319 SAVENTPKIQT-DRTKYETLLTQADLTWLEKLNAAKLAV-DTEIDSLDYSANLVGIS 376
OY 129 FVGLDIEWRPSFRKGVLPGRVATVOICVDSNYCDVMHIFHSGIPQSLQH-----LI 179
DB 377 F-----ALENGEAAYLPQLD--YLDA-----PKTLEKSTALAAIKPIL 413
OY 180 EDSTLVKVGIGIDGDSYKLFHDYGSVKIDVE-----DLSDLANOKIGCD 223
DB 414 ENPNHRIKQNIKFEDE-SIPARHGIELQGVFPTMLSTYLNSTGRHMDLAKRYLG-- 470
OY 224 KKGGLASLTETLVCKELKPNRIRLGMWEFPLSKQOLQYAAADAVASMLYKVL 278
DB 471 -----HETIAFESLAGKGSOL-TFNOIPL-EQATEYAADADVYMKLQOAL 515

RESULT 12
MSH6_ARATH
ID MSH6_ARATH STANDARD; PRT; 1324 AA.
AC 004716;
DT 15-DEC-1998 (Rel. 37, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA mismatch repair protein MSH6-1 (AtMSH6-1).
GN MSH6-1 OR AGAA.3 OR AT4G02070 OR T10M13.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. LANDSBERG ERECTA;
RA Till S., Granat S., Parnell L., Kaplan N., Hoffman J., Lodhi M.,
RA Johnson A.F., Dedhia N., Martienssen R., McCombe W.R.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA MEDLINE=99444907; PubMed=10517319;
RA Ade J., Beizille F., Philippe H., Dautriaux M.P.;
RT "Four mismatch repair paralogues coexist in Arabidopsis thaliana:
RT AtMSH2, AtMSH3, AtMSH6-1 and AtMSH6-2.";
RL Mol. Gen. genet. 262:239-249(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA MEDLINE=20083488; PubMed=10617198;
RA Mayet K.F.X., Scheller C., Wandute R., Murphy G., Volckaert G.,
RA Pohl T., Duesterhoeft A., Stiekema W., Entlan K.-D., Terry N.,
RA Harris B., Ansoorge W., Brandt P., Griwell L.A., Rieger M.,
RA Weichselgartner M., de Simone V., Obermayer B., Maché R., Mueller M.,
RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidheini T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Heineisel U., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.A., McCullagh B., Billham L., Robben J.,

```

RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
 RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
 RA Weizengger T., Bothé G., Ramsperger U., Hilbert H., Braun M.,
 RA Holzer E., Brandt A., Peters S., van Staveren M., Dirse W.,
 RA Moeljan P., Klein Lankhorst R., Rose M., Hauf J., Koelter P.,
 RA Berneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
 RA de Keyser A., Buysshert C., Gielen J., Villalroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIay K., Mayes R.,
 RA Pettitt A., Rajandream M.A., Lyne M., Benes V., Reichen M.,
 RA Borova D., Bloecker H., Scharfe M., Grimm M., Loehner T.-H.,
 RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Faltmann B., Granderath K., Dauner D., Hertzl A.,
 RA Neumann S., Argitlou A., Vitale D., Liguori R., Pitravden I.E.,
 RA Massenet O., Quigley F., Clabaud G., Muendlein A., Felber R.,
 RA Schabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 RA Chefor F., Cooke R., Berger C., Monfort A., Casachuberta E.,
 RA Gibbons T., Weber N., Vandenbol M., Barques M., Terol J., Torres A.,
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacón D., Jesse T.,
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bialke C.,
 RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
 RA Parnell L., Dehla N., Gnoj L., Schutz K., Huang E., Spiegel L.,
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Frieden J.,
 RA Storking T., Kallick J., Graves T., Harmon G., Edwards J.,
 RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
 RA Nelson J., Speith J., Ryan E., Andrews S., Geisel C., Layman D.,
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
 RA Antonola B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C.,
 RA Ma P., Zhong J., Preston R., Vill D., Shekher M., Mateo A., Shah R.,
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
 RA Granat S., Shobdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
 RA Chen E., Marra M., Martienssen R., McCombie W.R.;
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 thaliana." Nature 402:769-777(1999).
 RL Nature 402:769-777(1999).
 CC -1- FUNCTION: INVOLVED IN POST-REPLICATIVE DNA-MISMATCH REPAIR (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTS FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF001535; AAB57798.1; -
 DR EMBL: AJ245967; CAB53337.1; -
 DR EMBL: AF001308; AAC78699.1; -
 DR EMBL: AL161493; CAB80700.1; -
 DR InterPro: IPR000432; Muts_C.
 DR InterPro: IPR002863; Muts_N.
 DR InterPro: IPR002999; Muts_N.
 DR Pfam: PF00488; Muts_C.1.
 DR Pfam: PF01624; Muts_N.1.
 DR ProDom: PD001263; Muts_C.1.
 DR SMART: SM00534; Muts_C.1.
 DR SMART: SM00533; Muts_C.1.
 DR SMART: SM00333; TUDOR.1.
 DR PROSITE: PS00486; DNA_MISMATCH_REPAIR_2; 1.
 DR DNA repair: ATP-binding; DNA-binding.
 DR NP_BIND 1083 1090 ATP (POTENTIAL).
 FT NP_BIND 1083 1090 ADR -> GPKSKLL (IN REF. 1).
 FT CONFLICT 383 383 O -> OVRRAHNL (IN REF. 1).
 FT CONFLICT 852 852 P -> PGITNLHLITLILANTASHIISLP (IN
 FT REF. 1).
 SQ SEQUENCE 1324 AA: 146797 MW: 787A340272CE979C CRC64;

Query Match 6.0%; Score 89.5; DB 1; Length 1324;
 Best Local Similarity 23.2%; Predi No.15;
 Matches 68; Conservative 40; Mismatches 92; Indels 93; Gaps 18;
 QY 66 LPRSTISSTSYKRP-LSRCRANFPAMRGCR-----ILYSTATEVDKRAMOLIKVL- 118
 DB 766 LPYSLEFRKSLRLPMEERLIAMFSSIEASGRNGKVLVEDTA---RKQVDFISTLR 822
 QY 119 -----PKKR-----ESGIAVGDIDMRPFRKG-VL 145
 DB 823 GCTMAEACSSLAIRLKHDTSLRLHLTPGOSLPISSSIKFKAFDWEVHNAGRI 882
 QY 146 PKRVATVQICVDSNV---CDVMHIFHSIGIPQSH---LIEDSTLVKVGIGID----- 192
 DB 883 PHEGA-----DEYDCACTVEEFESSLKHLKEGRKLLGDASINVTYVGNDEYLLV 936
 QY 193 -GDSYKLFHDY-----GVS-----IKDVELSLDANQIGDKKGLASTITELY-- 236
 DB 937 ESLSGVPDHYELCSKKGSRYWPTIKL--LKLDSQAK--SEKESALKTSIQRLGR 992
 QY 237 -CKELKPNRIRLGNWFEYPLSKOQLOVATDAYSAMHLTVKLDIPDAVSGS 288
 DB 993 FCHQEK-----WQLVSAIHELVDLISLAFAS-DTEGYRCRP-VISGS 1035
 RESULT 13
 BLM DROME STANDARD; PRT; 1487 AA.
 ID BLM DROME
 AC 09VG18: 09Y062;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Bloom's syndrome protein homolog (EC 3.6.1.-) (Dmbim) (Mutagen-
 DE sensitive protein 309) (RecQ helicase homolog).
 GN mus109 or blm or CG6920.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxId=7227;
 RN [1]
 RP SEQUENCE FROM N.A., AND REPEATS.
 RC STRAIN=Canton-S;
 RX MEDLINE=99160561; PubMed=10049920;
 RA Kusano K., Berres M.E., Engels W.R.;
 RT "Evolution of the RECQ family of helicases: a Drosophila homolog,
 RT Dmbim, is similar to the human Bloom syndrome gene.";
 RL Genetics 151:1027-1039(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers J.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brocktein P., Brothier P.,
 RA Burks K.C., Busan P.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Folsler C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,

	FT	CONFLICT	594	594	D -> G (IN REF. 1).
	FT	CONFLICT	614	614	G -> A (IN REF. 1).
	FT	CONFICT	805	805	K -> E (IN REF. 1).
	FT	CONFLICT	1116	1116	E -> G (IN REF. 1).
	SQ	SEQUENCE	1487 AA:	166077 MW;	07361B8005E29432 CRC64;
		Query Match		6.0%;	Score 89.5; DB 1; Length 1487;
		Best Local Similarity	24.5%;	Pred.	No. 18;
		Matches	64; Conservative	42; Mismatches	94; Indels
					61; Gaps
					16;
	OY	60 NNIR-ROLP-SISITSSTYSKRFPLSRRCARNPAPRFGRIITYSKTATEVVKRAMOLIKV	117	:	: : :
	Bd	930 SNLRYRLPKRGVSITLDISRISK-----PQHSG-IITCLSRKECEDETSMKKMK-	980	:	:
	OY	118 LDRKDESGIAFVG---GLDIEMRPSEFKGYLPGKV---ATVOICVDNSNYCDVMHIFHS	169	'	:
	Bd	981 -----DGVRAVSYAHAGLTDTRESRCKDMVLTKGMRVICATAVRGMSIDKPDRVAFVLHY	1033	:	:
	OY	170 GIPOSLGHILDSLLKYKGIGIDGDVS*-LEHDYG--VISKDVEDLSDLANO----KIG	221	:	:
	Bd	1034 SLPSIEEGYOEAGR-----GRDGVDADCIITYNYSMDLRITKMID-SDKALQYNVKKH	1088	:	:
	OY	222 GDKMNGASLFETLV-CPELLKNPRIBLGNNFEVYLSCOOLO-----VKAT	266	:	:
	Bd	1089 VDNIIRTYGECENITDCRRAO-----QLDYGEHFTSQCLENRETACDNCRINKRAYAV	1143	:	:
	OY	.267 DAYASWHLKYVKLDLPDAVSG	287	::	: : ::
	Bd	1144 DALE-HARKAARA.VKDLCSG	1162	::	: : ::
		RESULT_14			
	ID	HK22_BRAVE	STANDARD;	PRT;	269 AA.
	AC	090481;			
	DT	15-JUL-1999 (Rel. 38,	Created)		
	DT	15-JUL-1999 (Rel. 38,	Last sequence update)		
	DT	15-JUL-1999 (Rel. 38,	Last annotation update)		
	DE	Homoebob protein NKX-2.2.			
	OS	NKX-2 OR NKX2.2 OR NKX-2.			
	GN	Brachydanio rerio (zebrafish) (Zebra danio).			
	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;			
	OC	Cypriniformes; Cyprindidae; Danio.			
	RN	NCBI_Taxid=7955;			
	XN	[1]			
	RP	SEQUENCE FROM N.A.			
	RX	MEDLINE-95324401; Pubmed-7600991.			
	RA	Barth K.A., Wilson S.W.:"Expression of zebrafish nkx2.2 is influenced by sonic hedgehog/vertebrate hedgehog-1 and demarcates a zone of neuronal differentiation in the embryonic forebrain."			
	RT	Development 121:1755-1768(1995).			
	RL	-I- SUBCELLULAR LOCATION: Nuclear (By similarity).			
	CC	-I- TISSUE SPECIFICITY: EXPRESSED IN A CONTINUOUS NARROW BAND OF CELLS ALONG A BOUNDARY ZONE DEMARCATING THE LOCATION AT WHICH TWO OF THE EARLIEST NUCLEI BELONGS TO THE NK-2 FAMILY OF HOMEBOX PROTEINS.			
	CC	-I- SIMILARITY: BELONGS TO THE NK-2 FAMILY OF HOMEBOX PROTEINS.			
	CC	-----			
	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isdb.ch/announce/or send an email to licensee@isb-sib.ch).			
	CC	-----			
	DR	EMBL; X85977; CAA59967.1; .			
	DR	HSSP; P22808; INK3.			
	DR	ZFIN; ZDB-GENE-980526-403; nkk2.2.			
	DR	InterPro; IPR001356; Homeobox.			
	DR	Pfam; PF00046; homeobox; 1.			
	DR	PRINTS; PR00024; HOMEOBOX.			

DR SMART: SM00389; HOX; 1.
 DR PROSITE: PS00027; HOMEBOX_1; 1.
 DR PROSITE: PS50071; HOMEBOX_2; 1.
 KM Homeobox: DNA-binding: Developmental protein; Nuclear protein.
 FT DNA_BIND 125 184 HOMEBOX
 SQ SEQUENCE 269 AA; 30306 MW; CF006285CDD33D63 CRC64;

Query Match 5.9%; Score 88; DB 1; Length 269;
 Best Local Similarity 22.2%; Pred. No. 2.4;
 Matches 45; Conservative 32; Mismatches 64; Indels 62; Gaps 8;

QY 8 DDATEEELALDAIEAS-VNFSRSSSSSAFTVQATTSVHGHEDDPNIPNNIRQL 66
 DB 69 DNPYT-RWLATTDSDIQLHGLSANSQDTSKSP-----EPSADESPNDKETS 116
 QY 67 PRSITSSTSYKRFPLSRNFRANFPMRFGRIILYKATAT-EVDKRAMQ----- 113
 DB 117 SNGSDSGKKRR-----RVLFKAQTYELERFRQRLSAPEREHL 158
 QY 114 -LIRKLDTK-----RDESGIAFVGLIDEMRPSFRKGVLPKGVATVQICVDSNYC 161
 DB 159 ASLRIRLPYQKIFQNHRYKMRARAEKMEVTHLPSPRAVAVPLVRDCKPCHTLKAQ 218
 QY 162 DVMIHFHSGIP-----QSLQHL 178
 DB 219 DLATFQAGIPFSAYSAOSLQHM 241

RESULT 15

PROA_VIBCH STANDARD; PRT; 416 AA.

AC Q9KPT9;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Gamma-glutamyl phosphate reductase (GPR) (EC 1.2.1.41) (Glutamate-5-semialdehyde dehydrogenase) (Glutamyl-gamma-semialdehyde dehydrogenase) (GSA dehydrogenase).
 DE PROA OR VC2273.
 GN *Vibrio cholerae*.
 OS *Vibrio cholerae*.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; *Vibrio*.
 OX NCBI_TaxId=666;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=EL TOR N16961 / SEROTYPE O1;
 RX MEDLINE=20406833; Pubmed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*."
 RL Nature 406:477-483(2000).
 CC -1- FUNCTION: CATALYZES THE NADPH DEPENDENT REDUCTION OF L-GAMMA-GLUTAMYL 5-PHOSPHATE INTO L-GLUTAMATE 5-SEMIALDEHYDE AND PHOSPHATE. THE PRODUCT SPONTANEOUSLY UNDERGOES CYCLIZATION TO FORM 1-PYRROLINE-5-CARBOXYLATE.
 CC -1- CATALYTIC ACTIVITY: L-glutamate 5-semialdehyde + phosphate + NADP(+) = L-gamma-glutamyl 5-phosphate + NADPH.
 CC -1- PATHWAY: PROLINE BIOSYNTHESIS PATHWAY; SECOND STEP.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE GAMMA-GLUTAMYL PHOSPHATE REDUCTASE FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.

CC EMBL: AE004299; AAF95417.1; ALT_INIT.
 DR TIGR: VC2273;
 DR InterPro: IPR002086; Aldehyde_dehydr.
 DR InterPro: IPR000965; GPR.
 DR Pfam: Pf00171; aldehyd; 1.
 DR PROSITE: PS01223; PROA; FALSE NEG.
 KM Oxidoreductase; Proline biosynthesis; NADP; Complete proteome.
 SQ SEQUENCE 416 AA; 44504 MW; 7ACAA58E537235563 CRC64;

Query Match 5.9%; Score 88; DB 1; Length 416;
 Best Local Similarity 21.7%; Pred. No. 4.3;
 Matches 69; Conservative 54; Mismatches 109; Indels 86; Gaps 19;

QY 7 IDDAFTEELALDAIEASVNFSRSSSSSAFTVQATTSVHGHEDDPNIPNNIR 63
 DB 59 LSDAMLDRLTL-----NESRLQATANDVRNVIKNDPV-GSEIDSRVLENGMSLAR 108
 QY 64 RQLPRSISSSTSYKRF-----LSRC-RANFPMRFGRIILYKATATEVDKRAMOLIK 116
 DB 109 RRVPLGV-GVYEARPVITIDIALCLTKGNAILNGKRETFPSN-----MELVK 158
 QY 117 VLDTRDESGIAFVGLD-IEWRPSFRKGVLPKGVATVQICVDSNYCDVMIHFHSGIPQ-- 173
 DB 159 VIOSALDKAGLPASAVQYIE-----KPDRELVTQLKMDDYVDMI-----IPRG 203
 QY 174 -SLGHLIEDSFLYKVGIGIDGDSVKLFHDYGVSLIKDYEDLSL-----ANOKIGDDK 225
 DB 204 AGLHKCKENSTVPVILG-----GFGISHIFVDESADLKSVAVENAKV--QRP 251
 QY 226 WGLASLETETVCKELKPNRIRL-----GNMEFY-PLSKOOLQYAA-TDAYA-----S 271
 DB 252 SACNALPDLVLVQALNPDLQKLAKLNGKVAFAEPRAKMLSSAELRDAQGDPDTE 311
 QY 272 WHLY-----KVLKQDPDAV 285
 DB 312 WLSYTLGVKVVQDVQEA 329

Search completed: October 31, 2002, 13:42:17
 Job time : 20 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 31, 2002, 13:40:24 ; Search time 18 Seconds

(without alignments)
1537.429 Million cell updates/sec

Title: US-09-896-186b-24

Perfect score: 1491

Sequence: 1 MSSSNWIDAFTEBELLAID.....YASWHLVYKLDLPDPAVSGS 288

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1439.5	96.5	313	2 T05256	hypothetical prote
2	339.5	22.8	1436	2 T14895	DNA helicase 1 - A
3	339	22.7	1401	2 T30247	Werner syndrome pr
4	330	22.1	1401	2 T17452	Werner syndrome pr
5	179.5	12.0	582	2 H96604	probable 3'-5' exo
6	158	10.6	123	2 E84752	hypothetical prote
7	136.5	9.2	910	2 S40930	hypothetical prote
8	127.5	8.6	239	2 A84777	hypothetical prote
9	112.5	7.5	217	2 T02548	hypothetical prote
10	101	6.8	876	2 T19246	hypothetical prote
11	97.5	6.5	445	2 E72674	hypothetical prote
12	96.5	6.5	488	2 F97039	hypothetical prote
13	95	6.4	508	1 A36121	hypothetical prote
14	95	6.4	4589	2 T14914	aromatase (EC 1.14
15	94.5	6.3	416	2 C71620	aromatase (EC 1.14
16	93.5	6.3	1525	1 GNMV55	protein with Egl-1
17	93	6.2	2224	1 KFHU5	genome polyprotein
18	93	6.2	3079	1 KEBY12	coagulation factor
19	92	6.2	237	2 H84732	probable GTPase-ac
20	92	6.2	758	2 T39210	hypothetical prote
21	91.5	6.1	473	2 S36553	3-isopropylmalate
22	91.5	6.1	781	2 T41551	L2 protein - human
23	91	6.1	503	2 S13912	hypothetical prote
24	90	6.0	472	2 T47436	aromatase (EC 1.14
25	89.5	6.0	930	2 E64098	protein kinase-11k
26	89.5	6.0	1324	2 T01508	DNA-directed DNA p
27	88	5.9	269	2 I50504	mismatch repair en
28	88	5.9	317	2 A97511	nk2.2 protein - ze
29	88	5.9	317	2 AH2729	hypothetical prote

30	88	5.9	417	2 E82095	gamma-glutamyl pho
31	88	5.9	709	2 S40926	hypothetical prote
32	88	5.9	784	2 C88558	protein ZK1098.3 l
33	87.5	5.9	334	2 T03556	probable D-ribose-
34	87.5	5.9	422	2 E96753	hypothetical prote
35	87.5	5.9	547	2 T40342	signal recognition
36	87.5	5.9	589	2 E88492	protein T07E3.1 l1
37	87	5.8	178	2 E72450	hypothetical prote
38	87	5.8	392	2 T44362	acetyl-CoA C-acety
39	87	5.8	579	2 T40371	methylintransferase
40	87	5.8	3137	1 A37797	collagen alpha 3(V
41	86.5	5.8	346	1 BYECPR	phosphate-repressi
42	86.5	5.8	360	2 T00882	hypothetical prote
43	86.5	5.8	470	2 T10193	glycerol-3-phospha
44	86.5	5.8	496	2 S61966	hypothetical prote
45	86.5	5.8	593	2 B89844	hypothetical prote

ALIGNMENTS

RESULT 1
T05256
hypothetical protein F18A5.260 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
C:Accession: T05256
R:Bevan, M.; Weber, N.; Grueninger, D.; Schmidheini, T.; Bancroft, I.; Mewes, H.W.; M
submitted to the Protein Sequence Database, February 1999
A:Reference number: Z15405
A:Accession: T05256
A:Molecule type: DNA
A:Residues: 1-313 <BEV>
A:Cross-references: EMBL:AL035528
A:Experimental source: cultivar Columbia; BAC clone F18A5.
C:Genetics:
A:Map position: 4
A:Introns: 86/1; 143/1; 186/3; 239/3
A:Note: F18A5.260

Query Match 96.5%; Score 1439.5; DB 2; Length 313;
Best Local Similarity 95.3%; Pred. No. 6e-117; 4; Indels 9; Gaps 1;
Matches 281; Conservative 1; Mismatches 4; Indels 9; Gaps 1;

QY	1	MSSSNWIDAFTEBELLAIDAIASYNFSRSSSSAAPTVOATTSVHGHEEDPNOIPN 60	
DB	1	MSSSNWIDAFTEBELLAIDAIASYNFSRSSSSAAPTVOATTSVHGHEEDPNOIPN 60	
QY	61	NIRQLPRSTSTSYKRPPLSCRARNPAMRFGGRILYSKTATEVDKRAMOLIVLDT 120	
DB	61	NIRQLPRSTSTSYKRPPLSCRARNPAMRFGGRILYSKTATEVDKRAMOLIVLDT 120	
QY	121	KRDESGIAFGLDIEMRSPRKGVATVOICVDSNCDVMHIFHSGIPSLDHLIE 180	
DB	121	KRDESGIAFGLDIEMRSPRKGVATVOICVDSNCDVMHIFHSGIPSLDHLIE 180	
QY	181	DSTLVVGICIDDSVKLFPHDYGVSIKIDVEDLSDLANOKIGDKKMGSLAFETLYCKEL 240	
DB	181	DSTLVVGICIDDSVKLFPHDYGVSIKIDVEDLSDLANOKIGDKKMGSLAFETLYCKEL 240	
QY	241	LKNNRIRLGNMERYPLSKOOLQYAATDAVASWHLVYVTTTKNHLTLINDLEAKIS 295	
DB	241	LKNNRIRLGNMERYPLSKOOLQYAATDAVASWHLVYVTTTKNHLTLINDLEAKIS 295	

RESULT 2
T14895
DNA helicase 1 - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T14895
R:Yan, H.; Chen, C.Y.; Kobayashi, R.; Newport, J.
Nature Genet. 19, 375-378, 1998

A:Title: Replication focus-forming activity 1 and the Werner syndrome gene product.

A:Reference number: Z18255; MUID:98361165

A:Accession: T14895

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1436 <YAN>

A:Cross-references: EMBL:AF067418; NID:g3420290; PID:g3420291; PIDN:AAC63512.1

C:Genetics:

A:Gene: FFA-1

Query Match 22.8%; Score 339.5; DB 2; Length 1436;

Best Local Similarity 35.0%; Pred. No. 8.6e-21;

Matches 82; Conservative 43; Mismatches 84; Indels 25; Gaps 7;

Db 61 NIPROLPRSTSTSYKRF---PLSRCRAR---NFPAMFGGRILYSKATEV-----D 108

Db 3 SLORRIPEMWSVYKQEDRIDDAKKSFKKNILEDNLPKMFNSIYYSNSCSSLSD 62

Db 109 KRAMOLIKVLTTRKDESGIAFYGLDIEMRPSFRKGVLPKRVATVOICVDSNYCDVMHIF- 167

Db 63 IRSSLL-----EDV---LGFDIEMPVYTKG-KTGKVALIOVCSEKCYLFPHISP 110

Db 168 HSGIPSOHLIEDSTLVYVGIGIDDSVYKLFYDGVSTIDVDLSDLANOKIGGDKWG 227

Db 111 MAFEPGRLKLEDESVRKVGIEGDQWKLMSDYELKIKGLTELEMANOKLRCKEKT 170

Db 226 LSLTLTVCKELLPNIRIRIGNMEFYPLSKOOLQYAAATDAYSMLYKVLKDL 281

Db 171 FNLIHLHFREQLYKRKSYRCNMDFILTEDOKLVAATDAYSMLYKVLKDL 224

RESULT 3
T30247
Werner syndrome protein type1 - mouse
N:Alternate names: Wrn type1 protein
C:Species: Mus musculus (house mouse)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 21-Jul-2000
C:Accession: T30247
R:Imamura, O.; Ichikawa, K.; Yamabe, Y.; Goto, M.; Sugawara, M.; Furuchi, Y.
Genomics 41, 298-300, 1997
A:Title: Cloning of a mouse homologue of the human Werner syndrome gene and assignment

A:Reference number: Z20785; MUID:97888537

A:Accession: T30247

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1401 <YMA>

A:Cross-references: EMBL:D86526; NID:g2130972; PIDN:BA02069.1; PID:g2130973

A:Experimental source: strain BALB/C; testis/spleen

C:Genetics:

A:Gene: WRN type1

A:Map position: 8A4

Query Match 22.7%; Score 339; DB 2; Length 1401;

Best Local Similarity 36.6%; Pred. No. 9.2e-21;

Matches 74; Conservative 42; Mismatches 72; Indels 14; Gaps 5;

Db 88 NFPAMFGGRILYSKATEYDKRAMOLIKVLTTRKDESGIAFYGLDIEMRPSFRKGVLP 147

Db 37 NLPLFEPGSIYVSASD---CSFLSEDISMRISDGV--VGFMEPPIYK-----PG 86

Db 148 K---VATVOICVDSNYCDVMHIFHSGI-POSTLOHLEDSTLVYVGIGIDDSVYKLFHDYG 203

Db 87 KRSRAVAVIQLCVSEKCYLFHISMSVFPQGLKMLLENSIKKAGVGIEGDQWKLRLRDF 146

Db 204 VSIRKVEDLSDLANOKIGGDKWGLASLTETLVCKELLPNIRIRIGNMEFYPLSKOOLQY 263

Db 147 VKLESVEELTDVANEKLCAEFTWSLGLVYKQLKDKSRCSNWSNFPITEDOKLY 206

Db 264 AATDAYASMLYKVLKDLDPDV 285

Db 207 AATDAYAGLIIYKLGKGLDITV 228

RESULT 4
T17452
Werner syndrome protein - mouse

N:Alternate names: Wrn protein

C:Species: Mus musculus (house mouse)

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T17452

R:Paepker, B.W.; Gayle, M.; Brady, W.; Swartz, A.; Gillett, L.A.; Allisch, R.S.; Mullig

A:Description: Genomic structure of the human Werner's gene and cloning of its mouse

A:Reference number: Z18794

A:Accession: T17452

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1401 <PAE>

A:Cross-references: EMBL:AF091215; NID:g3885837; PID:g3885838; PIDN:AAC78077.1

C:Genetics:

A:Gene: Wrn

Query Match 22.1%; Score 330; DB 2; Length 1401;

Best Local Similarity 36.0%; Pred. No. 5.5e-20;

Matches 72; Conservative 43; Mismatches 71; Indels 14; Gaps 5;

Db 88 NFPAMFGGRILYSKATEYDKRAMOLIKVLTTRKDESGIAFYGLDIEMRPSFRKGVLP 147

Db 37 NLPLFEPGSIYVSASD---CSFLSEDISMRISDGV--VGFMEPPIYK-----PG 86

Db 148 K---VATVOICVDSNYCDVMHIFHSGI-POSTLOHLEDSTLVYVGIGIDDSVYKLFHDYG 203

Db 87 KRSRAVAVIQLCVSEKCYLFHISMSVFPQGLKMLLENSIKKAGVGIEGDQWKLRLRDF 146

Db 204 VSIRKVEDLSDLANOKIGGDKWGLASLTETLVCKELLPNIRIRIGNMEFYPLSKOOLQY 263

Db 147 VKLESVEELTDVANEKLCAEFTWSLGLVYKQLKDKSRCSNWSNFPITEDOKLY 206

Db 264 AATDAYASMLYKVLKDLDPDV 283

Db 207 AATDAYAGLIIYKLGKGLDITV 226

RESULT 5
H96604
probable 3'-5' exonuclease [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: H96604

R:Rheologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Hultzar, L.

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Malt, R.; Matzla

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schmitz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: H96604

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-582 <STO>

A:Cross-references: GB:AE005173; NID:g11094727; PIDN:AAG29662.1; GSPDB:GN00141

C:Genetics:

A:Gene: F14G9.8

A:Map position: 1

Query Match 12.0%; Score 179.5; DB 2; Length 582;

Best Local Similarity 28.2%; Pred. No. 1.8e-07;

Matches 53; Conservative 36; Mismatches 66; Indels 33; Gaps 7;

Db 126 GIAFYGLDIEMRPSFRKGVLPKRVATVOICVDSN--YCDVMHIFHSGI-POSTLOHLE 180

Db 377 GCRVVGIDCEWKNYIKSKONKVSIMQIGSDTKIFILDLIKLYNDASEILDNCISHIQ 436

Db	43	DLPLLEFIHQIIVYSYDASDCSEFLBE-----DISMSLSDGGVYGFQDMENPPLVIRKGL-G	95
Qy	148	KVAIVQICVDSNYSYCDVMHIFHSIG-POSLOHLIEDSTLVKVGIGIDGDSYKLFPHDYVSI	206
		: : : : : : : : : : : : : : : : : : : :	
Db	96	KVALIQLCTVESKCYLFLFHVSSMSYFPGKLKMLLKNKAVKRVAGVIEEDQMKLLRDFPIKL	155
Qy	207	KDVEDLSLNAQKIGGDKKMGSLAETLVCKELKLRNRIRLGMWERYPLPSKOOLQYAAAT	266
		: : : : : : : : : : : : : : : : : : : :	
Db	156	KNFVELIDVARKKIKCTETWSLNSLVKHLGLQKOLLKDKSIKCSMWSKPEPLTEDOKLYAAT	215
Qy	267	DAYASWHLRYKVLKDLDPAY	285
		: : : : : : : : : : : : : : : : : : : :	
Db	216	DATAGFTIYRNLEITLDDTV	234

```

RESULT 2
US-09-127-670-6
: Sequence 6, Application US/09127670
: Patent No. 6328583
: GENERAL INFORMATION:
: APPLICANT: Massachussets Institute for Technology
: APPLICANT: Leonard P. Guarente
: APPLICANT: David A. Sinclair
: APPLICANT: David A. Lombard
: TITLE OF INVENTION: ASSAYS FOR COMPOUNDS WHICH EXTEND LIFE
: TITLE OF INVENTION: SPAN
: FILE REFERENCE: MIT-77209A
: CURRENT APPLICATION NUMBER: US/09/127,670
: CURRENT FILING DATE: 1998-07-31
: EARLIER APPLICATION NUMBER: US 60/054,629
: EARLIER FILING DATE: 1997-08-04
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 6
: LENGTH: 1401
: TYPE: .prt
: ORGANISM: Murline
: US-09-127-670-6

```

Query Match	22.78;	Score 339;	DB 4;	Length 1401;
Best Local Similarity	36.68;	Pred. NO. 2e-28;		
Matches	74;	Conservative	42;	Mismatches 72;
			Indels	14;
			Gaps	5

QY	88	NEPMARREGRLTYKTTAYEDVKRAMOLIKYVLDTRDSSGJAFVOLDIEMRSPFKGVLLP	147
Db	37	NLPFLPEFGSLVITYEASD-----CSFLSEDISMRLSGDY--VGFDMEPMPIYK----PG	86
QY	148	K---VATVQICVDSNVCYDVMIHFHSGI-PSLOHMLIEDSTLVYKGISIDGDSVLFEPDYG	203
Db	87	KRSRAVAVTQICVSENKCYLPHIISMSVFPQGLKMLLENKSIKKNGVOIEDGQMKLLNDFD	146
QY	204	VSIKDVEDLSLANOKTIGGDKKWKGLASLTETLVYCKELLKPNRIRLGNWEFYPISKQOLOY	263
Db	147	VKLESFVELTDPVANEKILKCAETWSLNGLVKHLVGLKQOLLKDKSIRCSNMSNPFLETDQKLY	206
QY	264	AATDAVASMHLKYVKLDLPDPAV	285
Db	207	AATDAVAGLIYQKLGINDGV	228

```

RESULT 3
US-08-781-891-206
? Sequence 206, Application US/08781891
? Patent No. 6090620
? GENERAL INFORMATION:
? APPLICANT: Fu, Ying-Hui
? APPLICANT: Yu, Chang-Ru
? APPLICANT: Oshima, Junko
? APPLICANT: Mulligan, John T.
? APPLICANT: Schellenberg, Gerald D.
? TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
? TITLE OF INVENTION: WERNER'S SYNDROME
? NUMBER OF SEQUENCES: 209

```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
City: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30D
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,891
FILING DATE: 27-DEC-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: NO. 6090620tendburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 240052.419
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 206:
SEQUENCE CHARACTERISTICS:
LENGTH: 1401 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-781-891-206

```

Query Match	22.1%;	Score 330;	DB 3;	Length 1401;
Best Local Similarity	36.0%;	Pred. NO. 2e-27;		
Matches	72;	Conservative	43;	Mismatches 71;
				Indels 14;
				Gaps 5;

[illegible]

```

1 RESULT 4
2 PCT-US94-00198-4
3 : Sequence 4 Application PC/TUS9400198 .
4 :
5 : GENERAL INFORMATION:
6 :
7 : APPLICANT: Schering Corp.
8 :
9 : TITLE OF INVENTION: RAS Associated GAP Proteins
10 :
11 : NUMBER OF SEQUENCES: 6
12 :
13 : CORRESPONDENCE ADDRESS:
14 :
15 : ADDRESSEE: Schering Corp.
16 :
17 : STREET: 1 Giraldo Farms
18 :
19 : CITY: Madison
20 :
21 : STATE: New Jersey
22 :
23 : COUNTRY: USA
24 :
25 : ZIP: 94304-1104
26 :
27 : COMPUTER READABLE FORM:
28 :
29 : MEDIUM TYPE: Floppy disk
30 :
31 : COMPUTER: Macintosh
32 :
33 : OPERATING SYSTEM: 6.0.8
34 :
35 : SOFTWARE: Microsoft Word 5.1a
36 :
37 : CURRENT APPLICATION DATA:
38 :
39 : APPLICATION NUMBER: PCT/US94/00198

```

```

: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/004,824
: FILING DATE: 15-JAN-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Lunn, Paul G.
: REGISTRATION NUMBER: 32,743
: REFERENCE/DOCKET NUMBER: DX0352 PCT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (201)822-7255
: TELEFAX: (201)822-7039
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3079 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: ORIGINAL SOURCE:
: ORGANISM: Saccharomyces cerevisiae
PCT-US94-00198-4
```

Query Match 5.2%; Score 93; DB 5; Length 3079;

Best Local Similarity 19.3%; Pred. No. 1.7; Indels 84; Gaps 10;

Matches 57; Conservative 43; Mismatches 112;

```

QY 9 DAFTEELALDAIEASYNFRSSSSSAAPVQATTSVHGHEEDPNQIPNNIRQLPR 68
DB 1976 DDFYKTFLLIDVGLQLOPKMEFSN-----ETPIYREMDYDPELEYEMNNHAFR 2028
QY 69 SITSSSY-----KRPPLSRCRRANFPMPREGRLYSKTATENDKRAMOLIK-- 116
DB 2029 NIETSTAVSPSVHESTSESEIPITLTMSNF-----SDRHIDIDTVAVKFLQIY 2077
QY 117 -----VLTKRDESGIAFVGLDIEMRPSFRKGVLPKQATVQICDSVYCDVMHI 166
DB 2078 ARWTKRHCLIDCTERDEG---GLDMKRKFLVWGLLP-EVAP-KNCIGCYFVNNET 2131
QY 167 F-----HSGIPQSLQHLIEDSTLVKVGIGIDGDSVKLFHDYGSIKRVED 211
DB 2132 FMNYGKCLDKDNVYSSKPHFINSNSDEGLMK-SVGTGGGLAVLQDIRSLMDITL 2190
QY 212 LSLANOKIGGDKKGLASITFLVCKELKPNRIRLGNMEFYPLSKOOLQVATD 267
DB 2191 YDEKRRN-----FTPVSLKIGDIYFOVLHETPROXYKIRD 2224
```

RESULT 5

US-09-722-139-2

```

: Sequence 2, Application US/09722139
: Patent No. 6355471
: GENERAL INFORMATION:
: APPLICANT: Beraud, Christophe
: APPLICANT: Freedman, Richard
: TITLE OF INVENTION: No. 6355471el motor proteins and methods for
: FILE OF INVENTION: their use
: FILE REFERENCE: 1055
: CURRENT APPLICATION NUMBER: US/09/722,139
: CURRENT FILING DATE: 2000-11-24
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 2
: LENGTH: 1375
: TYPE: PRT
: ORGANISM: Human
US-09-722-139-2
```

Query Match 5.9%; Score 87.5; DB 4; Length 1375;

Best Local Similarity 19.0%; Pred. No. 2;

Matches 58; Conservative 52; Mismatches 116; Indels 79; Gaps 13;

QY 13 EEELALDAIEASYNFRSSSSSAAPVQATTSVHGHEEDPNQIPNNIRQLPRSTS 72

```

DB 935 KEEOLA-----OYQANANOLOKLOATFEFTANLAROE-----KVRKKEKILES 979
QY 73 STYKRPPLSRCRRANFPMPREGRLYSKTATENDKRAMOLIKVLDTRKDESGIAFVG 132
DB 980 REKOQREALERALAR--LEKRHSALDRHSLTGLTEIEFQKOKLASLNSGSRREGSLQ-ASL 1036
QY 133 DIEMRPSFRK-----GVLPGRVATVQICVDSNYCDVMH 165
DB 1037 EAE-QEALKEQDRLLEYEIQLKQIYEVGVQKDHGTLLEGKVASSSLPVSAEKSHLVP 1095
QY 166 I-----FHSGIPQSLQHLIED-STLVVVGIGIDGDSVKLFHDYGVSTIKDVEDLSLANOKI 220
DB 1096 LMDARINAYIEEVEVORLQDLNHRVISEGCSSTADTK-----DNEKLNHGTI 1142
QY 221 GGDKKWGLASITFLVCKELKPNRIRLGNMEFYPLSKOOLQVATDAYASWHLVYVKMD 280
DB 1143 QRLKLY---ELCRDLICVLMPEPDAACAN---HPLLQDLVLSLD---WK-----TE 1187
QY 281 LPDAV 285
DB 1188 IPLDV 1192
```

RESULT 6

US-07-989-845-2

```

: Sequence 2, Application US/07989845
: Patent No. 5304472
: GENERAL INFORMATION:
: APPLICANT: Bass, Steven
: APPLICANT: Swartz, James
: TITLE OF INVENTION: METHOD OF CONTROLLING POLYPEPTIDE
: TITLE OF INVENTION: PRODUCTION IN BACTERIAL CELLS
: NUMBER OF SEQUENCES: 31
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genentech, Inc.
: STREET: 460 Point San Bruno Blvd
: CITY: South San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94080-4990
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: patin (Genentech)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/989,845
: FILING DATE: 19921120
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Hasak, Janet E.
: REGISTRATION NUMBER: 28,616
: REFERENCE/DOCKET NUMBER: 752
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415/225-1896
: TELEFAX: 415/952-9881
: TELEX: 910/371-7168
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 348 amino acids
: TYPE: AMINO ACID
: TOPOLOGY: linear
US-07-989-845-2
```

Query Match 5.8%; Score 86.5; DB 1; Length 348;

Best Local Similarity 20.9%; Pred. No. 0.3;

Matches 54; Conservative 34; Mismatches 73; Indels 97; Gaps 12;

QY 86 ARNFPMPREGRLYSKTATV-----DKRAM-----OLIKVLDTRKDESG 126

Db 107 AVNIPGLKSGELVLDGKTLGDIYLGKIKKWDDEAIAKLNPLGKLPSONIAV--RRADGSG 165
QY 127 IAFV-----GLDIEMRPSRKGV-----LPGVVAIVQICVDS 158
Db 166 TSVFTSYLAKVNEKNNVGTSTYKWPILGIGKNGDIAIAVQRIPGAIGVEYA-- 222
QY 159 NYCDVWHIFHSGIPQSLQHLIEDSTLVKVGIGIDGSVKLFHDYGVSIKVEDLS----- 213
Db 223 -YAKONNLATYTKL-----ISADGKPVSTPEENFANNAKAGADMSTFQ 264
QY 214 DLANOKIGDKKWLASLETTLVCKELKPNR-----IRLGNMEFYPLSKQ--OLOYAATD 267
Db 265 DLTNOK--GEDAMPITSTFILHDKOKKPEQGEVLKFFDMAVYKTGAKQANDLDYAS-- 320
QY 268 AYASWHLKYVLKDLDPAY 285
Db 321 -----LPDSV 325

RESULT 7

PCT-US93-11298-2
Sequence 2, Application PC/TUS9311298
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
TITLE OF INVENTION: METHOD OF CONTROLLING POLYPEPTIDE PRODUCTION IN
TITLE OF INVENTION: BACTERIA
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 Inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: palin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/11298
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Hasak, Janet E.
REGISTRATION NUMBER: 28, 616
REFERENCE/DOCKET NUMBER: 752
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1896
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 348 amino acids
TYPE: amino acid
TOPOLOGY: linear
PCT-US93-11298-2

Query Match 5.8%; Score 86.5; DB 5; Length 348;
Best Local Similarity 20.9%; Pred. No. 0.3;
Matches 54; Conservative 34; Mismatches 73; Indels 97; Gaps 12;

QY 86 ARNFPAMRFGRLYSKTAIEV-----DKRAM-----OLIKVLDTKRDESG 126
Db 107 AVNIPGLKSGELVLDGKTLGDIYLGKIKKWDDEAIAKLNPLGKLPSONIAV--RRADGSG 165
QY 127 IAFV-----GLDIEMRPSRKGV-----LPGVVAIVQICVDS 158
Db 166 TSVFTSYLAKVNEKNNVGTSTYKWPILGIGKNGDIAIAVQRIPGAIGVEYA-- 222

QY 159 NYCDVWHIFHSGIPQSLQHLIEDSTLVKVGIGIDGSVKLFHDYGVSIKVEDLS----- 213
Db 223 -YAKONNLATYTKL-----ISADGKPVSTPEENFANNAKAGADMSTFQ 264
QY 214 DLANOKIGDKKWLASLETTLVCKELKPNR-----IRLGNMEFYPLSKQ--OLOYAATD 267
Db 265 DLTNOK--GEDAMPITSTFILHDKOKKPEQGEVLKFFDMAVYKTGAKQANDLDYAS-- 320
QY 268 AYASWHLKYVLKDLDPAY 285
Db 321 -----LPDSV 325

RESULT 8

US-08-907-166-12
Sequence 12, Application US/08907166
Patent No. 5948666
GENERAL INFORMATION:
APPLICANT: Callen, Walter
TITLE OF INVENTION: ISOLATION AND IDENTIFICATION OF NOVEL POLYMERASES
FILE REFERENCE: 09010/027001
CURRENT APPLICATION NUMBER: US/08/907,166
CURRENT FILING DATE: 1997-08-06
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows version 3.0
SEQ ID NO: 12
LENGTH: 574
TYPE: PRT
ORGANISM: Aquifex pyrophilus
US-08-907-166-12

Query Match 5.4%; Score 80; DB 2; Length 574;
Best Local Similarity 21.9%; Pred. No. 3.5;
Matches 33; Conservative 30; Mismatches 48; Indels 40; Gaps 6;

QY 170 GIPOSLOHLIEDSTLVKVGIGIDGSVKLF-----HDYGV--SIKVEDLSLANOK- 219
Db 11 GLKKAIRL-ENSPYLYLDTETTGDRIRLVQIGDEENTYVIDLYEIODIEPLKLINEG 69
QY 220 -IGDKKWL-----ASLETTLVCKELKPNRIL-----CN 250
Db 70 YGHNILKFDLKYLRGIPPSATFFDMIASYLLGYERHSLNIHVSNLGYSMDKSYOTSD 129
QY 251 WEFPYLSKQOLOYAATDAVASWHLKYVLKDL 281
Db 130 WGNASVLSDAQLKTAANDVYLVRLFPKMRDM 160

RESULT 9

US-08-845-258-34
Sequence 34, Application US/08845258
Patent No. 6183976
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Lodes, Michael J.
APPLICANT: Houghton, Raymond
APPLICANT: Sleath, Paul R.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS


```

; LENGTH: 431 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
;
US-08-723-142A-34

```

Query Match	5.3%	Score 79.5	DB 4	Length 431
Best Local Similarity	22.4%	Pred. No. 2.5		
Matches 66	Conservative 35	Mismatches 106	Indels 87	Gaps 17

```

QY 9 DAFEEELALDALEASYNFSRSSSSSSSAAPTQOATTVSHGHEEDNOJPNNRQJPR 68
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 33 DNFILDNPSAFLIELRVSYNSNEFOYOS-----PONTNNEESSTPE 73

QY 69 S-----ITSSTSYKREPLSRCRANFPAMREGILYSKTATEYDKRAMOLIKVLTDRD 123
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 74 .SNITVVHSDVMKRF--NCKNKR-----KSLSTHSLTEND-----ILKGRRI 113

QY 124 ESGI--AFVGDIDIMRPSFRK---VLPGKATVQIOVDSXCYDVMH-IFHSIGIPOLQH 177
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 114 ELAKYCIIMGAGITASDLNLGLCFLFISDKOST--NVC--NYFEDMHESHILDTQASD 168

QY 178 LIEDS-----TLVKVG--IGIDGDSYKLFPHDYGSIKIDVEDLSLDAOKIGGRKW 226
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 170 CVSDGDGADIDISNEDFYQDGINSVADSETCMANSQVYNNPNEVNSER-----F 221

QY 227 G-LASL--TELVKCKELLKPRIRIKLGNWEPFL-----SKOQLOYAATDAYASW 272
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 222 GKLSLVSTTPPLCRICL-----CGSDDPPLVTPPCNKGSLNYVHLECLRTW 269

```

RESULT 12
US-08-895-601-6
; Sequence 6, Application US/08895601

```

1  GENERAL INFORMATION:
2
3  APPLICANT:      Beer-Romero, Peggy
4
5  APPLICANT:      Strack, Peter J.
6
7  APPLICANT:      Glass, Susan J.
8
9  APPLICANT:      Rolfe, Mark
10
11 TITLE OF INVENTION:  REGULATION OF KAPPA B (1KB) DEGRADATION
12
13 TITLE OF INVENTION:  AND METHODS AND REAGENTS RELATED THERETO
14
15 NUMBER OF SEQUENCES: 16

```

ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30

APPLICATION NUMBER: US/08/895,601
FILING DATE: 16-JUL-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 56,709
REFERENCE/DOCKET NUMBER: MIV-096.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000

```

? TELEFAX: 617-893-7000
? INFORMATION FOR SEQ ID NO: 6
? SEQUENCE CHARACTERISTICS:
?   LENGTH: 927 amino acids
?   type: amino acid
?   TOPOLOGY: linear
? MOLECULE TYPE: protein
?
JS-08-895-601-6

```

Query Match	5.3%	Score 79	DB 3	Length 927
Best Local Similarity	20.9%	Pred. No. 9.5		
Matches 63	Conservative 37	Mismatches 82	Indels 120	Gaps 19

```

Qy      22 IEASINFSRSSSSSSAAPTVOATTSVHGH-EDPNOIPNNIRQLPSITSSTSYKRF 80
      ::: |::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      408 VQATVETSQSSSAGPQSQASTSDSGQVTPSEIEQGF---LPRG----- 455

```

Qy	81	LSRCRANRFPAMRFGGRLLY	-----SKTATEVDRKAMOLIV	-----LDTRKD	-----123
		1	11	11	11
Db	454	WEVRHAP	NGREFLDHNTKTTTWEDPR	-----LKPIAHLRKTSLDTSNDLGPL	502
Qy	124	-----ESGIAFVGLDI	---EMR-PSFRKVLDPKVAATVOICVDSNYCDVMHIF	-----167	
		1	1	1	1
Db	503	PCGWERTHTGRIIFYIHNHNKRPQWEDPRLENAIIGPA	-----VPYSRQYKRKKYEFR	558	
Qy	168	-----HSGIPQSLQ	-----HLIEQS	-----TLVKGIGIGDSGVPLPHDYGSIK	207
		1	1	1	1
Db	559	KLKKONDIPIPKFEKKLRATVLEDSYRRIMGVKRADEFLKARLWIEFDEKGL	---DYG---	613	
Qy	208	DVEDLSDLANOKIGDKKRWGLASTETPLVCKELKPNRIRLGNWMEFPLSKOOLQYATD	-----267		
		1	1	1	1
Db	614	-----GVAREMWF	-----LISEMENP	-----YVGL	-----FEVSATD 611
Qy	268	AY 269			
Db	642	NY 643			

RESULT 13
US-07-640-029-2
; Sequence 2, Application US/07640029

GENERAL INFORMATION:
APPLICANT: Kiefer, Michael C.
APPLICANT: Valenzuela, Pablo D.T.
APPLICANT: Barr, Philip J.
TITLE OF INVENTION: Expression and Use of Human Fibroblast
TITLE OF INVENTION: Growth Factor Receptor
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:

1 ADDRESSSEE: Chicon Horton Corporation
2 STREET: 4560 Horton Street
3 CITY: Emeryville
4 STATE: California
5 COUNTRY: USA
6 ZIP: 94608
7
8 COMPUTER READABLE FORM:
9 MEDIUM TYPE: Floppy disk
10 COMPUTER: IBM PC compatible
11 OPERATING SYSTEM: PC-DOS/MS-DOS
12 SOFTWARE: Patentin Release #1.0; Version #1.2
13 CURRENT APPLICATION DATA:
14 APPLICATION NUMBER: US/07/6440, 029
15 FILING DATE: 19910111

```

1  ATTORNEY/AGENT INFORMATION:
2
3  NAME: McClung, Barbara G.
4
5  REGISTRATION NUMBER: 33,113
6
7  REFERENCE/DOCKET NUMBER: CH-165
8
9  TELECOMMUNICATION INFORMATION:
10
11  TELEPHONE: 510-601-2708
12
13  TELEFAX: 510-655-3542
14
15  INFORMATION FOR SEQ ID NO: 2:
16
17  SEQUENCE CHARACTERISTICS:
18
19  LENGTH: 817 amino acids
20
21  TYPE: AMINO ACID
22
23  STRANDEDNESS: single
24
25  TOPOLOGY: linear
26
27  MOLECULE TYPE: peptide
28
29  US-07-640-029-2

```

US-07-640-029-2

Query Match 5.3%: Score 78.5; DB 1; Length 817;
Best Local Similarity 18.2%: Pred. No. 8.8;
Matches 56; Conservative 46; Mismatches 99; Indels 107; Gaps 11;

QY 1 MSSSWIIDA-----FTEELLALDAIEASYNFSRSSSSSSAAPT-----VOATT 46
Db 61 VQSIWMLWDGVOQLAESNFRITRTGEEVQDSVPADSGIYACVTSPPSGSDTTYSVNVSD 120
QY 47 SVHGHEEDPNOIPNNIRQLPRSTSTSYKRFPLSRCRANFPAMRFGRIILYSKATE 106
Db 121 ALPSEDDDDDDSSSEK-----ETDNTKPRMPVAP-----WTSSEK 160
QY 107 VDKR-AMOLIKVLDTRKDESG-----IAFGLDIEMRPSRKGYLPKQVATVQICVS- 158
Db 161 MEKRLHAVPAKATVFKCPSSTGTPNPTLRMLKNGKEFKPDRIIGYKVRATWSTIMDSV 220
QY 159 -----NY-----CDVMH-----IFHSGIP----- 172
Db 221 VPSDKGNTTCIYENEGYSINHTYQLDVERSPHRIILQAGLPANKTVALGSNVERMCKVY 280
QY 173 -----OSLOHLIEDST-----LVKVGIGIDGSVKLFHDYGVSIKDVEDLS 213
Db 281 SDPQHIOMLKHIENYSGSKIGPDNLPHYQILKTAGVNTTDKEMEVLHLRNVSFEDEAGET 340
QY 214 DLANKIG 221
Db 341 CLAGNSIG 348

RESULT 14
US-07-997-133-1
Sequence 1, Application US/07997133
Patent No. 528885
GENERAL INFORMATION:
APPLICANT: Bergonzoni, Laura
APPLICANT: Mazze, Guy
APPLICANT: Isacchi, Antonella
APPLICANT: Roncucci, Romeo
APPLICANT: Sarnientos, Paolo
TITLE OF INVENTION: Extracellular Form of the Human
TITLE OF INVENTION: Fibroblast Growth Factor Receptor
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/997,133
FILING DATE: 28-DEC-1992
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/642,755
FILING DATE: 18-JAN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 528885man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 769-226-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 24885 OPAT UR
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 822 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-997-133-1

Query Match 5.3%: Score 78.5; DB 1; Length 822;
Best Local Similarity 18.2%: Pred. No. 8.9;
Matches 56; Conservative 46; Mismatches 99; Indels 107; Gaps 11;

QY 1 MSSSWIIDA-----FTEELLALDAIEASYNFSRSSSSSSAAPT-----VOATT 46
Db 61 VQSIWMLWDGVOQLAETNTRITGEEVQDSVPADSGIYACVTSPPSGSDTTYSVNVSD 120
QY 47 SVHGHEEDPNOIPNNIRQLPRSTSTSYKRFPLSRCRANFPAMRFGRIILYSKATE 106
Db 121 ALPSEDDDDDDSSSEK-----ETDNTKPRMPVAP-----WTSSEK 160
QY 107 VDKR-AMOLIKVLDTRKDESG-----IAFGLDIEMRPSRKGYLPKQVATVQICVS- 158
Db 161 MEKRLHAVPAKATVFKCPSSTGTPNPTLRMLKNGKEFKPDRIIGYKVRATWSTIMDSV 220
QY 159 -----NY-----CDVMH-----IFHSGIP----- 172
Db 221 VPSDKGNTTCIYENEGYSINHTYQLDVERSPHRIILQAGLPANKTVALGSNVERMCKVY 280
QY 173 -----OSLOHLIEDST-----LVKVGIGIDGSVKLFHDYGVSIKDVEDLS 213
Db 281 SDPQHIOMLKHIENYSGSKIGPDNLPHYQILKTAGVNTTDKEMEVLHLRNVSFEDEAGET 340
QY 214 DLANKIG 221
Db 341 CLAGNSIG 348

RESULT 15
US-07-921-807B-4
Sequence 4, Application US/07921807B
Patent No. 5474914
GENERAL INFORMATION:
APPLICANT: SPAETE, RICHARD
TITLE OF INVENTION: METHOD OF INCREASING EXPRESSION
TITLE OF INVENTION: OF VIRAL PROTEINS
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: CHIRON CORPORATION
STREET: 4560 Horton Street - R440
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/921,807B
FILING DATE: 29-SEP-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MCCUNG, BARBARA G.
REGISTRATION NUMBER: 33,113
REFERENCE/DOCKET NUMBER: 0209,001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2708
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 822 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-921-807B-4

Query Match 5.3%; Score 78.5; DB 1; Length 822;
Best Local Similarity 18.2%; Pred. No. 8.9;
Matches 56; Conservative 46; Mismatches 99; Indels 107; Gaps 11;

QY 1 MSSNNIADDA-----FTEELAIADAIASINFSRSSSSSAAPT---VOATT 46
: | | | : | | | : | | | : | | | :
Db 61 VOSINMLRDGVQLAESNRRTITGEVEVQDSVPADSGLYACVTSSPSGSDTTYFSVNVSD 120
: | | | : | | | : | | | : | | | :
QY 47 SVHGHEEDPQIINNIRROLPRSTSTSYKRPPLSRCRARNPAMRFGGRILYSKTATE 106
: | | | : | | | : | | | : | | | :
Db 121 ALPSSSEDDDDDDSSSEER---ETDWTKPNRMPVAP-----YWTSPK 160
: | | | : | | | : | | | : | | | :
QY 107 VDKR--AMOLIKVLDTRKDESG-----IAFVGIDIEWRPSFRKGVLPKVAIVQICVDS- 158
: | | | : | | | : | | | : | | | :
Db 161 MEKKLHVPAAKTIVKFCPSGTPNPTLRMLKNGKEKPDHRIGYKVRATWSIIMDSV 220
: | | | : | | | : | | | : | | | :
QY 159 -----NY-----CDVMH-----IFHSGIP----- 172
: | | | : | | | : | | | : | | | :
Db 221 VPSDKGNVTCIVENEGSINHYYQLDVERSPPHPILOAGLPANKTVALGSNVEFMCKVY 280
: | | | : | | | : | | | : | | | :
QY 173 -----OSLOHLIEDST-----LVKVGIGIDGSVKLFHDYGVSIKDVEDLS 213
: | | | : | | | : | | | : | | | :
Db 281 SDPQPHIQMLKHLIEVNGSKIGPDNLPYVQILKTAGVNTTDPKMEVHLHNRNVSPEDAGEYT 340
: | | | : | | | : | | | : | | | :
QY 214 DLANKIG 221
: | | | : | | | : | | | : | | | :
Db 341 CLAGNSIG 348

Search completed: October 31, 2002, 13:43:34
Job time : 19 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 31, 2002, 13:36:19 ; Search time 34 Seconds
(without alignments)
940,860 Million cell updates/sec

Title: US-09-896-186b-24

Perfect score: 1491

Sequence: 1 MSSSMWIDAFTEELLAIID.....YASWHLKVKLDLPDAVSGS 288

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802:*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1491	100.0	288	21	AA652858
2	1491	100.0	298	21	AA652857
3	1491	100.0	298	21	AA652856
4	1028	68.9	197	21	AA652855
5	996	66.8	197	21	AA652854
6	896	60.1	177	21	AA652853
7	352	23.6	1432	18	AA114517
8	339	22.7	1401	19	AAW59454
9	339	22.1	1401	20	AAW97841
10	330	22.1	1401	18	AA114519
11	229.5	15.4	346	22	ABB71575

12	212.5	14.3	583	22	ABB62653	Drosophila melanog
13	178.5	12.0	244	22	ABB80638	Environmental stre
14	151	10.1	127	22	ABB11264	Human secreted pro
15	151	10.1	147	20	AA929336	Human secreted pro
16	151	10.1	147	22	AA939087	Human secreted pro
17	151	10.1	466	22	AA939116	Human protein sequ
18	105	7.0	625	22	ABB58154	Drosophila melanog
19	103	6.9	220	21	AA604684	Arabidopsis thalia
20	96	6.4	706	16	AA75647	Thermophilic bacte
21	95	6.4	161	21	AA628363	Arabidopsis thalia
22	95	6.4	167	21	AA628362	Arabidopsis thalia
23	94.5	6.3	416	21	AA818185	Novel human diagno
24	93	6.2	325	22	ABG17672	Human Factor V. H
25	93	6.2	2224	17	AAW04254	Human lipoprotein
26	93	6.2	2224	20	AA49564	Arabidopsis thalia
27	91	6.1	451	21	AA618364	Arabidopsis thalia
28	91	6.1	477	21	AA618363	Arabidopsis thalia
29	89.5	6.0	346	15	AA60663	psts variant. Esc
30	89.5	6.0	930	22	AA635540	Haemophilus influe
31	89.5	6.0	1078	22	ABB61252	Drosophila melanog
32	89.5	6.0	1487	22	ABB62792	Drosophila melanog
33	89	6.0	991	21	AA83171	Cell wall protein
34	89	6.0	991	21	AA770120	Staph. epidermidis
35	88.5	5.9	3079	15	AA659926	GAP protein Iraz.
36	88	5.9	506	22	ABB64451	Drosophila melanog
37	87.5	5.9	642	22	ABB61316	Drosophila melanog
38	87.5	5.9	2354	22	ABB60511	psts variant. Esc
39	86.5	5.8	346	15	AA60664	psts variant. Esc
40	86.5	5.8	346	15	AA60664	psts variant. Esc
41	86.5	5.8	346	15	AA60664	psts variant. Esc
42	86.5	5.8	346	15	AA60664	psts variant. Esc
43	86.5	5.8	346	15	AA60664	psts variant. Esc
44	86.5	5.8	346	15	AA60664	psts variant. Esc
45	86.5	5.8	346	15	AA60664	psts variant. Esc

ALIGNMENTS

RESULT 1

ID AAG52858 standard; Protein; 288 AA.

XX AC AAG52858;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 67237.

XX KW Protein identification; signal transduction pathway; metabolic pathway;

XX KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EPI033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 9905-0121825.

XX PR 05-MAR-1999; 9905-0123180.

XX PR 09-MAR-1999; 9905-0123548.

XX PR 23-MAR-1999; 9905-0125788.

XX PR 25-MAR-1999; 9905-0126264.

XX PR 29-MAR-1999; 9905-0126785.

XX PR 01-APR-1999; 9905-0127462.

XX PR 06-APR-1999; 9905-0128234.

XX PR 08-APR-1999; 9905-0128714.

XX PR 16-APR-1999; 9905-0129845.

XX PR 19-APR-1999; 9905-0130077.

XX PR 21-APR-1999; 9905-0130449.

PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134258.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143452.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.

PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 1491; DB 21; Length 288;
Best Local Similarity 100.0%; Pred. No. 4,6e-151;
Matches 288; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSNMWIDAFTEEELAIIDAIEASYNFSRSSSSSAAPTVOATTSVHGHEDPNQIPN 60
DB 1 MSSNMWIDAFTEEELAIIDAIEASYNFSRSSSSSAAPTVOATTSVHGHEDPNQIPN 60
QY 61 NTRQLPRSTSTSTSKRRPLSKRCRANRPNMFGGRILYSKATTEVDKRAMOLIVLDT 120
DB 61 NTRQLPRSTSTSTSKRRPLSKRCRANRPNMFGGRILYSKATTEVDKRAMOLIVLDT 120
QY 121 KRDESGIAVGLDIEWRPFGRKGLVPGKATVQICVDSNCDVMHIFHSGIPQSLQHLIE 180
DB 121 KRDESGIAVGLDIEWRPFGRKGLVPGKATVQICVDSNCDVMHIFHSGIPQSLQHLIE 180
QY 181 DSTLVKVGIGIDGDSVKLEFHDYGVSIKDYEDLSDLANOKRIGGDKKMGSLSTETLVCKEL 240
DB 181 DSTLVKVGIGIDGDSVKLEFHDYGVSIKDYEDLSDLANOKRIGGDKKMGSLSTETLVCKEL 240
QY 241 LKPNRRLGNWFEYPLSKQOLQYATDAYASWHLIYVKLDPDAVSGS 288
DB 241 LKPNRRLGNWFEYPLSKQOLQYATDAYASWHLIYVKLDPDAVSGS 288

RESULT 2

AAG52857 standard; Protein; 298 AA.

AAG52857;

18-OCT-2000 (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO: 67236.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

OS EPI033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-012180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129645.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145813.
PR 27-JUL-1999; 99US-0145818.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.

PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 1491; DB 21; Length 298;
Best Local Similarity 100.0%; Pred. No. 4, 9e-151;

Matches 288; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSSNWIDDAFTEEBELAIIDAIEASYNFSRSSSSSSSAFTVOATTSVHGHEEDPNOIPN 60
DB 11 MSSSSNWIDDAFTEEBELAIIDAIEASYNFSRSSSSSSSAFTVOATTSVHGHEEDPNOIPN 70
QY 61 NRRQLPRSTTSSTSKRPFPLSRCRANRPFAMRFGGRILYKATVEYDKRAMOLIVLDT 120
DB 71 NRRQLPRSTTSSTSKRPFPLSRCRANRPFAMRFGGRILYKATVEYDKRAMOLIVLDT 130
QY 121 KRDESGIAFYGLDIEWRPFRKGVLPKATVQICVDSNYCDVMHIFHSGIPQSLQHLIE 180
DB 131 KRDESGIAFYGLDIEWRPFRKGVLPKATVQICVDSNYCDVMHIFHSGIPQSLQHLIE 190
QY 181 DSTLVKVGIGIDGDSVKLFHDYGVSTIKDVEDLSDLANOKTGGKKNGLASTTETLVCKEL 240
DB 191 DSTLVKVGIGIDGDSVKLFHDYGVSTIKDVEDLSDLANOKTGGKKNGLASTTETLVCKEL 250
QY 241 LKPNRIRLGNMEFFPLSKOOLQYAATDAYSMHLVYKLDLPDAVSGS 288
DB 251 LKPNRIRLGNMEFFPLSKOOLQYAATDAYSMHLVYKLDLPDAVSGS 298

RESULT 3

ID AAG24977 standard; protein; 288 AA.

XX AAG24977;

XX 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 28856.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

OS Arabidopsis thaliana.
XX

PN EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0122180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0131048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144532.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0146738.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150684.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152263.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.

PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161992.
PR 28-OCT-1999; 9905-0161993.
PR 29-OCT-1999; 9905-0162142.

Query Match 96.1%; Score 1433; DB 21; Length 288;
Best Local Similarity 96.5%; Pred. No. 7.6e-145;

Matches 278; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 MSSSNWIDAFTEEEELALDAIEASYNFSRSSSSSAFTVATTSVHGHEEDPNQIPN 60
DB 1 MSSSNWIDAFTEEEELALDAIEASYNFSRSSSSSAFTVATTSVHGHEEDPNQIPN 60
QY 61 NIKROLPRSTTSSTYKRFPFLSCRRANFPAMRFGGRILYKSTATEVDKRAMOLIKVLD 120
DB 61 NIKROLPRSTTSSTYKRFPFLSCRRANFPAMRFGGRILYKSTATEVDKRAMOLIKVLD 120
QY 121 KRDESGIAFYGLDIEMRPPSRKGVLPCKVATVQICVDNSNCDVMHFFHSGIPQSLQHLIE 180
DB 121 KRDESGIAFYGLDIEMRPPSRKGVLPCKVATVQICVDNSNCDVMHFFHSGIPQSLQHLIE 180
QY 181 DSTLVKVGIGIDGSYKLFPHDYGVSIKDYEDLSDLANOKIGDKKMGCLASTLETIVCKEL 240
DB 181 DSTLVKVGIGIDGSYKLFPHDYGVSIKDYEDLSDLANOKIGDKKMGCLASTLETIVCKEL 240
QY 241 LKPNRIRLGNMEFYPLSKOOLQYAAATDAYASWHLKYVKLDPDAVSDS 288
DB 241 LKPNRIRLGNMEFYPLSKOOLQYAAATDAYASWHLKYVKLDPDAVSDS 288

RESULT 4
AA652859
ID AA652859 standard; Protein; 197 AA.
AC AA652859;
XX
XX
BT 18-OCT-2000 (first entry)

XX XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 67238.
XX XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX XX
OS Arabidopsis thaliana.
PN EP1033405-A2.
PD 06-SEP-2000.
XX XX
PF 25-FEB-2000; 2000EP-0301439.
XX XX
PR 25-FEB-1999; 9905-0121825.
PR 05-MAR-1999; 9905-0123180.
PR 09-MAR-1999; 9905-0123548.
PR 23-MAR-1999; 9905-0125788.
PR 25-MAR-1999; 9905-0126264.
PR 29-MAR-1999; 9905-0126785.
PR 01-APR-1999; 9905-0127462.
PR 06-APR-1999; 9905-0128234.
PR 08-APR-1999; 9905-0128714.
PR 16-APR-1999; 9905-0129845.
PR 19-APR-1999; 9905-0130077.
PR 21-APR-1999; 9905-0130077.
PR 23-APR-1999; 9905-0130510.
PR 23-APR-1999; 9905-0130891.
PR 28-APR-1999; 9905-0131449.
PR 30-APR-1999; 9905-0132048.
PR 30-APR-1999; 9905-0132407.
PR 04-MAY-1999; 9905-0132484.
PR 05-MAY-1999; 9905-0132485.
PR 06-MAY-1999; 9905-0132486.
PR 06-MAY-1999; 9905-0132487.
PR 07-MAY-1999; 9905-0132863.
PR 11-MAY-1999; 9905-0134256.
PR 14-MAY-1999; 9905-0134218.
PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134370.
PR 18-MAY-1999; 9905-0134768.
PR 19-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 21-MAY-1999; 9905-0135353.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 17-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139492.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139763.

PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.

PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156559.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match

Best Local Similarity 100.0%; Pred. No. 1,le-101;

Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 MRGGILVSKTATEVDKRAMOLIKVLDTRKRDSSGIAFPGLDIEMPRSRKGLPKKVAFT 151
DB 1 MRGGILVSKTATEVDKRAMOLIKVLDTRKRDSSGIAFPGLDIEMPRSRKGLPKKVAFT 60
QY 152 VOICVSNYCDVWHIFHSGIPQSLQHLIEDSTLVKVGIGIDGSVLTFHDYGVSIKDVED 211
DB 61 VOICVSNYCDVWHIFHSGIPQSLQHLIEDSTLVKVGIGIDGSVLTFHDYGVSIKDVED 120
QY 212 LSDLANOKTIGDCKKWLASITETLVCKELLKPNRIRLGNWFEYPLSKOOLQYAAATDAYAS 271
DB 121 LSDLANOKTIGDCKKWLASITETLVCKELLKPNRIRLGNWFEYPLSKOOLQYAAATDAYAS 180
QY 272 WHLYKVLKDLPPAVSGS 288
DB 181 WHLYKVLKDLPPAVSGS 197

RESULT 5
AAG24978

ID AAG24978 standard; Protein: 197 AA.
XX
AC AAG24978;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 28857.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
PN EPI03405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 200DEP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.

PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151348.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 66.8%; Score 996; DB 21; Length 197;
Best Local Similarity 97.0%; Pred. No. 2.9e-98;
Matches 191; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 92 MRFGRIYSKTTEVDKRAMQILKVLDTFRDSSGIAFVGLDEMRPSFKGVLPGVAT 151
Db 1 MRFGRIYSKTTEVDKRAMQILKVLDTFRDSSGIAFVGLDEMRPSFKGVLPGVAT 60
OY 152 VOICVDNNTCDVMHIFHSGIPQSLQHLIEDSTLVKVGIGIDGDSVKLFDHYGVSIXKVED 211
Db 61 VOICVDNNTCDVMHIFHSGIPQSLQHLIEDSTLVKVGIGIDGDSVKLFDHYGVSIXKVED 120
OY 212 LSLDAQKIGGDKKMGSLATETLVCKELKPKRIRIGNMEFPLSKQOLQYATDAYAS 271
Db 121 LSLDAQKIGGDKKMGSLATETLVCKELKPKRIRIGNMEFPLSKQOLQYATDAYAS 180
OY 272 WHLYKVLKDLPDVAVSGS 288
|||||

Db 181 WHLYKVLKDLPDVAVSDS 197
RESULT 6
AAG24979
ID AAG24979 standard; Protein; 177 AA.
XX AAG24979;
XX
XX 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 28858.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
PN EP1033405-A2.
PN
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138545.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.

Db 121 TESTVCKELLKPNKIRLGNWFFYPLSKOOLQYATDAYSAMHLYKVLKDLDPDAVS DS 177

RESULT 7
AAV14517
ID AAV14517 standard; Protein; 1432 AA.

AC AAV14517;
DT 31-AUG-1999 (first entry)
XX

DE Human WRN gene product.

KM Human; WRN; Werner's syndrome; detection; diagnosis; autosomal;
KM recessive disorder; phenotype.

OS Homo sapiens.

PN WO9724435-A1.

PD 10-JUL-1997.

PF 30-DEC-1996; 96MO-US20785.

PR 12-APR-1996; 96US-0632175.

PR 29-DEC-1995; 95US-0009409.

PR 29-DEC-1995; 95US-0580539.

PR 30-JAN-1996; 96US-0010835.

PR 30-JAN-1996; 96US-0594242.

PA (DARW-) DARWIN MOLECULAR CORP.

PA (OSHI/) OSHIMA J.

PI Fu Y, Mulligan J, Oshima J, Schellenberg GD, Yu C;

PI WPI; 1997-363671/33.

DR N-PSDB; AAX83001.

DR N-PSDB; AAX83001.

PT Isolated nucleic acid molecule encoding the WRN gene product -

PT useful for detection and treatment of Werner's syndrome, and related

PT diseases

XX Claim 10; Fig 2B; 153pp; English.

XX This sequence represents the human WRN gene product which is associated

CC with Werner's syndrome. The products can be used for the detection and

CC treatment of Werner's syndrome (WS), an autosomal recessive disorder

CC with a complex phenotype, as well as related diseases.

XX Sequence 1432 AA;

XX

XX

XX

XX

XX

XX

RESULT 8
AAW59454
ID AAW59454 standard; Protein; 1401 AA.

AC AAW59454;

DT 28-AUG-1998 (first entry)

DE Mouse WRN helicase protein.

KM Werner's syndrome; WRN; helicase; murine; diagnosis.

OS Mus sp.

PN JP10146188-A.

PD 02-JUN-1998.

PF 15-NOV-1996; 96JP-0304721.

PR 15-NOV-1996; 96JP-0304721.

PA (EIJT-) EIJIN KENKYUSHO KK.

DR WPI; 1998-369863/32.

DR N-PSDB; AAV35114, AAV35115.

PT New mouse gene encoding protein - used for diagnosis of human

PT Werner's syndrome

XX Claim 1; Page 18-22; 30pp; Japanese.

XX This sequence represents a novel mouse WRN helicase protein which is

CC associated with Werner's syndrome. The mouse WRN gene is useful for

CC the diagnosis of human Werner's syndrome.

XX Sequence 1401 AA;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

RESULT 9
AAW97841
ID AAW97841 standard; Protein; 1401 AA.

AC AAW97841;

DT 07-JUN-1999 (first entry)

DE Murine mWRN polypeptide (WRN homologue).

KM mWRN gene; WRN gene; homologue; Werner's syndrome; mouse;

KM life span; ageing.

OS Mus sp.

PN

PD

PF

PR

PA

DR

DR

PT

PT

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX MO9906543-A1.
 PN 11-FEB-1999.
 XX
 PF 03-AUG-1998; 98MO-US16081.
 XX
 PR 04-AUG-1997; 97US-0054629.
 XX
 PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
 XX
 PI Guarente LP, Lombard DB, Sinclair DA;
 XX
 DR MPI; 1999-153773/13.
 DR N-PSDB; AAX24302.
 XX
 PT Identifying agents that inhibit replication and accumulation of DNA
 PT circles - as potential agents for increasing the life span of
 PT animals
 XX
 PS Claim 13; Page 69-72; 78pp; English.

CC This is the amino acid sequence of a polypeptide encoded by mWRN
 CC see AAX24302), the murine homologue of the human WRN gene associated
 CC with Werner's syndrome. A claimed assay for identifying a compound
 CC which extends life span comprises administering a compound to be
 CC tested to a mammal with a suppressed level of mWRN, and identifying
 CC slowing of at least one of the normal phenotypes of ageing in the
 CC mammal. A compound identified by this method inhibits the
 CC replication and/or accumulation of rDNA circles in a cell. The
 CC assay can be performed using a mWRN knockout mouse (also claimed).
 CC mWRN nucleic acids can be used to produce mWRN polypeptides and as
 CC diagnostic probes and primers. mWRN polypeptides can be used to
 CC raise antibodies (for inhibiting activity of mWRN or for determining
 CC it in tissues). The assay is based on the observation that in yeast
 CC cells, accumulation of ribosomal circular DNA is responsible for
 CC age-related enlargement and fragmentation of the nucleolus. A
 CC mutation in the yeast WRN homologue SGS1 causes premature ageing,
 CC suggesting a common ageing mechanism in all eukaryotes. Insight
 CC into the ageing process in model systems can provide insight into
 CC ageing in humans.

XX Sequence 1401 AA:

Query Match 22.7%; Score 339; DB 20; Length 1401;
 Best Local Similarity 36.6%; Pred. No. 1.5e-26;
 Matches 74; Conservative 42; Mismatches 72; Indels 14; Gaps 5;

QY 88 NPFAMRFGRLSKTATEVDKRAMOLIKVLDTRKRDSEGIAPFGLDIEMRPFRRKGVLP 147
 DB 37 NLFPLFPGSIVSYEASD---CSFLSEDISMRISDGDV--VGFDMEWPPIYK---PG 86
 QY 148 K---VATVQICVDNSYCDVNHIFHSGI-POSTLOHLEDSTLVKVGIGIDDSVKLFHDYG 203
 DB 87 KRSRVAVIQCSESKCYLFIHISMSVFPQGLMKLENNKSIKRAVGIEEDQMKLLRDPD 146
 QY 204 VSKIDVEDSLDNOKIGCKWGLASLTETLVCKELLRNRIPLGNWMEYPPYLSKOOLY 263
 DB 147 VKLESFVELTDVANEKIKCAETWSLNGLVKLVGKQLKDKSIRCSNWSNFPLEDOCKLY 206
 QY 264 AATDAYASWHLVYKVLKDPDAV 285
 DB 207 AATDAYAGLITTYOKLGNLGDV 228

RESULT 10
 ID AAY14519 standard; Protein; 1401 AA.
 AC AAY14519;
 XX
 DT 31-AUG-1999 (first entry)
 XX

DE Mouse WRN gene product.
 XX
 KW Human: WRN; Werner's syndrome; detection; diagnosis; autosomal;
 KW recessive disorder; phenotype.
 XX
 OS Mus musculus.
 XX
 PN MO9724435-A1.
 XX
 PD 10-JUL-1997.
 XX
 PF 30-DEC-1996; 96MO-US20785.
 XX
 PR 12-APR-1996; 96US-0632175.
 PR 29-DEC-1995; 95US-0009409.
 PR 29-DEC-1995; 95US-0580539.
 PR 30-JAN-1996; 96US-0010835.
 PR 30-JAN-1996; 96US-0594242.

PA (DARN-) DARWIN MOLECULAR CORP.
 PA (OSHI/) OSHIMA J.
 XX

PI Fu Y, Mulligan J, Oshima J, Schellenberg GD, Yu C;
 XX
 DR MPI; 1997-363671/33.
 DR N-PSDB; AAX83004.
 XX

PT Isolated nucleic acid molecule encoding the WRN gene product
 PT useful for detection and treatment of Werner's syndrome, and related
 PT diseases

PS Claim 10; Fig 6; 153pp; English.

CC This sequence represents the mouse WRN gene product. The corresponding
 CC human gene product (AAY14517) is associated with Werner's syndrome. The
 CC products can be used for the detection and treatment of Werner's
 CC syndrome (WS), an autosomal recessive disorder with a complex phenotype,
 CC as well as related diseases.

XX Sequence 1401 AA:

Query Match 22.1%; Score 330; DB 18; Length 1401;
 Best Local Similarity 36.0%; Pred. No. 1.4e-25;
 Matches 72; Conservative 43; Mismatches 71; Indels 14; Gaps 5;

QY 88 NPFAMRFGRLSKTATEVDKRAMOLIKVLDTRKRDSEGIAPFGLDIEMRPFRRKGVLP 147
 DB 37 NLFPLFPGSIVSYEASD---CSFLSEDISMRISDGDV--VGFDMEWPPIYK---PG 86
 QY 148 K---VATVQICVDNSYCDVNHIFHSGI-POSTLOHLEDSTLVKVGIGIDDSVKLFHDYG 203
 DB 87 KRSRVAVIQCSESKCYLFIHISMSVFPQGLMKLENNKSIKRAVGIEEDQMKLLRDPD 146
 QY 204 VSKIDVEDSLDNOKIGCKWGLASLTETLVCKELLRNRIPLGNWMEYPPYLSKOOLY 263
 DB 147 VKLESFVELTDVANEKIKCAETWSLNGLVKLVGKQLKDKSIRCSNWSNFPLEDOCKLY 206
 QY 264 AATDAYASWHLVYKVLKDPD 283
 DB 207 AATDAYAGLITTYOKLGNLGD 226

RESULT 11
 ID ABB71575 standard; Protein; 346 AA.
 AC ABB71575;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide seq ID NO 41517.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;

XX Environmental stress resistance; salt; heat; desert; transgenic plant.
KW Saleola komarovii.
OS WO200106006-A1.
PN 25-JAN-2001.
XX 19-JUL-2000; 2000WO-JP04862.
XX 19-JUL-1999; 99JP-0235910.
PR 24-MAR-2000; 2000JP-0085377.
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
PA Yamada A, Ozeki Y, Satou T;
XX WPI; 2001-147355/15.
DR N-PSDB; AAF74217.
XX Screening method to obtain DNA encoding environmental stress resistance
PT factor, useful for producing transgenic plants resistant to
PT environmental stress -
XX Claim 64; Page 159-160; 167pp; Japanese.
XX Polynucleotide sequences AAF74187 - AAF74218 encode proteins
CC AAB80608 - AAB80639, which impart environmental stress resistance. The
CC invention relates to a method for identifying DNA encoding proteins
CC imparting environmental stress resistance. The method comprises inserting
CC cDNA from a library originating in a salt-resistant organism into a host
CC cell, culturing the transformants under conditions in which the
CC untransformed host does not grow well, and selecting for viable clones.
CC The method is useful for obtaining DNA encoding environmental stress
CC resistance factors. The DNA encoding proteins conferring environmental
CC stress resistance, can be used in the production of plants resistant to
CC environmental stress, which can be cultivated in unfavourable
CC environments such as deserts, salt damaged ground, cold regions and the
CC oceans. They can be used for increasing the area of land covered by green
CC plants, and desert greening and afforestation, in order to counter the
CC effects of the increase in atmospheric carbon dioxide concentration. PCR
CC primers AAF74219 and AAF74220 are used in an example illustrating the
CC method of the invention.
CC
SQ Sequence 244 AA:
Query Match 12.0%; Score 178.5; DB 22; Length 244;
Best Local Similarity 32.0%; Pred. No. 1.7e-10;
Matches 48; Conservative 28; Mismatches 57; Indels 17; Gaps 7;
QY 130 VGLDIEMRPSFRKGLPGKATVQICVDSNYCDVMHIFHS-GIPQSHLIEDSTLVKVG 188
DB 55 IGLGVQMKPS-----STSAATIQLSIDKR-CLIFQLSHSPALIPATLRDLDDRYTFFG 107
QY 189 I--GIDGDSVKLFHDYGVSIKIDVEDSLANQKIGDKKMGSLTETLV--CKELKPN 244
DB 108 VHNQGRARDLLOGSHNE---LDVNNLVDLAEENGHILKMSMEDMADVLGFC-GVHKPR 162
QY 245 RIRLGMMERPLSKOOLQVATATAYASWML 274
DB 163 KVMLSGMDQYCLSNDOYVACVADAVYSLRL 192
RESULT 14
ABBI1264
ID ABBI1264 standard; peptide; 127 AA.
AC ABBI1264;
XX 11-JAN-2002 (first entry)
DT Human secreted protein homologue, SEQ ID NO:1634.
DE

XX Human: cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolytic; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antisthmatic; antiarthritis; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnery; antitumor.
XX Homo sapiens.
OS
XX WO200157188-A2.
PN
XX 09-AUG-2001.
PD
XX 05-FEB-2001; 2001WO-US03800.
PF
XX 03-FEB-2000; 2000US-0496914.
PR
XX 27-APR-2000; 2000US-0560875.
XX
XX (HYSE-) HYSEQ INC.
PA
PI Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-457740/49.
DR N-PSDB; ABA08508.
XX Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer -
XX
XX Claim 20; Page 161; 1963pp; English.
XX Sequences ABBI0981-ABBI1230 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides

CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a novel human
CC polypeptide of the invention.

XX Sequence 127 AA;

Query Match 10.1%; Score 151; DB 22; Length 127;
Best Local Similarity 35.0%; Pred. No. 5.4e-08;
Matches 41; Conservative 20; Mismatches 48; Indels 8; Gaps 2;

OY 171 IPQSLQHLIEDSTLVKVGIGIDGDSVKLFPHDYGVSIRKVEDLSLA-----NOKIGGDKK 225

DB 3 LPRLLDILADGTLTKVGVCSGSDASKLQDYGIVRGCDLRYLAMRQRNNLLCNG--- 59

OY 226 WGLASLTETLVCKELKPNRIRLGNMEFFPLSKQQLQYAATDAYASMHLYKVLKDL 282

DB 60 LSLKSLAETVYNFPPLDKSLLRCSNMDAETLTEDOVITYAARDQAISVALFHLILGYP 116

RESULT 15

AA29336 ID AAY29336 standard; Protein: 147 AA.

XX AC AAY29336;

XX DT 29-SEP-1999 (first entry)

XX DE Human secreted protein clone cs756_2 alternate reading frame protein.

KW Human; secreted protein; nutrition; cytokine; cell proliferation;
KW differentiation; immune stimulating; vaccine; suppression; gene therapy;
KW haematopoiesis regulation; tissue growth; activin; inhibin; cadherin;
KW chemotactic; chemokinetic; haemostatic; thrombolytic; anti-inflammatory;
KW tumour invasion suppressor; tumour inhibition.

XX OS Homo sapiens.

XX PN W09937674-A1.

XX PD 29-JUL-1999.

XX PF 21-JAN-1999; 99WC-US01404.

XX PR 20-JAN-1999; 99US-0235609.

XX PR 22-JAN-1998; 98US-0072134.

XX PA (GEMV) GENETICS INST INC.

PI Agostino MJ, Clark HF, Collins-Racie LA, Fechtel K;
PI Jacobs K, Lavallic ER, McCoy JM, Merberg D, Steininger RJ;
PI Treacy M, Wong GG;

XX WPI: 1999-458682/38.

XX DR N-PSDB; AAX90441.

XX PT New polynucleotides encoding secreted human proteins derived from,
XX e.g. fetal brain potentially used as immunostimulators.

XX PS Disclosure; Page 133-134; 139pp; English.

XX The present sequence represents a human secreted protein. Human secreted
CC protein polynucleotides and proteins are predicted to have biological
CC activities which would make them suitable for treating, preventing or
CC ameliorating medical conditions in humans and animals, although no
CC supporting data is given. Suggested activities include nutritional
CC activity, cytokine and cell proliferation/differentiation activity,
CC immune stimulating (e.g. as vaccines) or suppressing activity,
CC haematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cadherin/tumour invasion suppressor activity, and tumour
CC inhibition activity. The polynucleotides are also stated to be useful
CC for gene therapy.

XX Sequence 147 AA;

Query Match 10.1%; Score 151; DB 20; Length 147;
Best Local Similarity 35.0%; Pred. No. 6.8e-08;
Matches 41; Conservative 20; Mismatches 48; Indels 8; Gaps 2;

OY 171 IPQSLQHLIEDSTLVKVGIGIDGDSVKLFPHDYGVSIRKVEDLSLA-----NOKIGGDKK 225

DB 23 LPRLLDILADGTLTKVGVCSGSDASKLQDYGIVRGCDLRYLAMRQRNNLLCNG--- 79

OY 226 WGLASLTETLVCKELKPNRIRLGNMEFFPLSKQQLQYAATDAYASMHLYKVLKDL 282

DB 80 LSLKSLAETVYNFPPLDKSLLRCSNMDAETLTEDOVITYAARDQAISVALFHLILGYP 136

Search completed: October 31, 2002, 13:41:54
Job Time : 38 secs

